Simple Dynamics for Majority Consensus*

L. Becchetti[1], A. Clementi[2], E. Natale[1], F. Pasquale[1], R. Silvestri[1], and L. Trevisan[3]

[1]*Sapienza* Università di Roma, `becchett@dis.uniroma1.it`, `natale@di.uniroma1.it`, `pasquale@di.uniroma1.it`, `silvestri@di.uniroma1.it`
[2]Università *Tor Vergata* di Roma, `clementi@mat.uniroma2.it`
[3] Stanford University, `trevisan@stanford.edu`

October 14, 2013

Abstract

We study a *Majority Consensus* process in which each of n anonymous agents of a communication network supports an initial opinion (a color chosen from a finite set $[k]$) and, at every time step, he can revise his color according to a random sample of neighbors. It is assumed that the initial color configuration has a sufficiently large *bias* s towards a fixed majority color, that is, the number of nodes supporting the majority color exceeds the number of nodes supporting any other color by an additive factor s. The goal (of the agents) is to let the process converge to the *stable* configuration where all nodes support the majority color. We consider a basic model in which the network is a clique and the update rule (called here the *3-majority dynamics*) of the process is that each agent looks at the colors of three random neighbors and then applies the majority rule (breaking ties uniformly).

We prove that the process converges in time $\mathcal{O}\left(\min\{k,(n/\log n)^{1/3}\}\,\log n\right)$ with high probability, provided that $s \geqslant c\sqrt{\min\{2k,(n/\log n)^{1/3}\}\, n\log n}$. Departing significantly from the previous analysis, our proof technique also yields a $\mathrm{polylog}(n)$ bound on the convergence time whenever the initial number of nodes supporting the majority color is larger than $n/\mathrm{polylog}(n)$ and $s \geqslant \sqrt{n\,\mathrm{polylog}(n)}$, *no matter how large k is.* We then prove that our upper bound above is tight as long as $k \leqslant (n/\log n)^{1/4}$. This fact implies an exponential time-gap between the majority-consensus process and the *median* process studied in [6].

A natural question is whether looking at more (than three) random neighbors can significantly speed up the process. We provide a negative answer to this question: in particular, we show that samples of polylogarithmic size can speed up the process by a polylogarithmic factor only.

*Partially supported by Italian MIUR-PRIN 2010-11 Project *ARS TechnoMedia* and the EU FET Project *MULTIPLEX* 317532.

1 Introduction

We consider a communication network in which each of n anonymous nodes supports an initial opinion (a color chosen from a finite set $[k]$). In the *Majority-Consensus* problem, it is assumed that the initial color configuration has a sufficiently large *bias* s towards a fixed majority color $m \in [k]$ - that is, the number c_m of nodes supporting the majority color (in short, the *initial majority size*) exceeds the number c_j of nodes supporting any other color j by an additive factor s - and the goal is to design an efficient fully-distributed protocol that lets the network converge to the *majority consensus*, i.e., to the monochromatic configuration in which all nodes support the majority color.

Reaching majority consensus in a distributed system is a fundamental problem arising from several areas such as Distributed Computing [6, 18], Communication Networks [19], and Social Networks [5, 16, 15]. Inspired by some recent works analyzing simple updating-rules (called *dynamics*) for this problem [1, 6], we study a discrete-time, synchronous process in which, at every time step, each of the n anonymous nodes can revise his color according to a random sample of neighbors.

We consider one of the simplest models, in which the network is a clique, and the update rule, called here 3-*majority dynamics*, is that each node samples at random three neighbors, and picks the majority color among them (breaking ties uniformly). Let us remark that looking at less than three random neighbors would yield a coloring process that converges to a non-majority color with constant probability even for $k = 2$ and large initial bias (i.e. $s = \Theta(n)$).

In [6], a tight analysis of a 3-neighbor dynamics for the *median* problem on the clique has been presented: the goal here is to converge to a stable configuration where all nodes support a value which is a "good" approximation of the *median* of the initial color configuration. It turns out that, in the binary case (i.e $k = 2$), the median problem is equivalent to the majority consensus one and the 3-input dynamics for the median is equivalent to the 3-majority dynamics: as a result, they obtain, for any bias $s \geqslant c\sqrt{n \log n}$ for some constant $c > 0$, an optimal bound $\Theta(\log n)$ on the convergence time of the 3-majority dynamics for the binary case of the problem considered in this paper.

However, for any $k \geqslant 3$, it is easy to see that the two problems above are different from each other (the median may be very different from the majority value) and, thus, the two dynamics are different from each other as well. Moreover, the analysis in [6] - strongly based on the properties of the median function - cannot be adapted to bound the convergence time of the majority process. The existence of dynamics reaching majority consensus for the general case is left as an important open issue in [2, 6, 1].

Our contribution. We present a new analysis of the 3-majority dynamics in the general case (i.e. for any $k \in [n]$). Our analysis shows that, with high probability (in short, *w.h.p.*), the process converges to majority consensus within time $O\left(\min\{k, (n/\log n)^{1/3}\} \log n\right)$, provided that the initial bias is $s \geqslant c\sqrt{\min\{2k, (n/\log n)^{1/3}\}\, n \log n}$, for some constant $c > 0$. Observe that the required minimal bias can be negligible, i.e., $s = \Theta(n^\epsilon)$, for some constant $\epsilon < 1$. A further analysis on the required initial bias is given in Section 5 and in Appendix A.3.

Our proof technique is accurate enough to get another interesting form of the above upper bound that does not depend on k. Indeed, when the initial majority size c_m is larger than $n/\lambda(n)$ for any function $\lambda(n)$ such that $3 \leqslant \lambda(n) < \sqrt{n}$ and $s \geqslant \sqrt{\lambda(n)\, n \log n}$, then the process w.h.p. converges in time $O\left(\lambda(n) \log n\right)$, no matter how large k is. Hence, when

$c_m \geqslant n/\mathrm{polylog}(n)$ and $s \geqslant \sqrt{n\,\mathrm{polylog}(n)}$, the convergence time is polylogaritmic.

We then show that our upper bound is tight for a wide range of the input parameters. When $k \leqslant (n/\log n)^{1/4}$, we in fact prove a lower bound $\Omega(k \log n)$ on the convergence time of the 3-majority dynamics starting from some configurations with bias $s \leqslant (n/k)^{1-\epsilon}$, for arbitrarily small constant $\epsilon > 0$. Observe that this range largely includes the initial bias required by our upper bound when $k \leqslant (n/\log n)^{1/4}$. So, the *linear-in-k* dependence of the convergence time cannot be removed for a wide range of parameter k.

Our analysis also provides a clear picture of the 3-majority dynamic process. Informally speaking, the larger the initial value of c_m is (w.r.t. n), the smaller the required initial bias s and the faster the convergence time are. On the other hand, our lower-bound argument shows, as a by-product, that the initial majority size c_m needs $\Omega(k \log n)$ rounds just to increase from $n/k + o(n/k)$ to $2\,n/k$.

We then prove a general negative result: in the considered distributed model, there is no dynamics with at most 3 inputs (but the majority one) that w.h.p. converges to majority consensus starting from any initial bias s such that $s = o(n)$. In other words, not only there is no hope to find a 3-input dynamics that is asymptotically faster than $k \log n$ but the 3-majority dynamics is the only one working for the majority consensus, no matter in how much time. Rather interestingly, by comparing the $O(\log n)$ bound for the median process [6] to our negative results for the majority process on the same distributed model, we get an exponential time-gap between the the task of computing the median and that of computing the majority (consensus) (this happens for instance when $k = n^a$, for any constant $0 < a < 1/4$).

A natural question arising from our results is whether a (slightly) larger random sample of neighbors might lead to a significant speed-up of the convergence time to majority consensus. We provide a negative answer to this question. We consider the generalization of the 3-majority dynamics, the h-majority one, where every node, at every time step, updates his color according to the majority of the colors supported by an arbitrary number h of random neighbors. We prove a lower bound $\Omega\left(k/h^2\right)$ on the convergence time of the h-majority, for integers k and h such that $k/h = O\left(n^{1/4-\epsilon}\right)$, where ϵ is an arbitrarily-small positive constant. We emphasize that scalable and efficient protocols must yield low communication complexity and small node congestion at every time step. These properties are guaranteed by the h-majority dynamics only when h is small, say $h = O(\mathrm{polylog}(n))$: in this case, our lower bound says that the resulting speed up is only polylogarithmic with respect to the 3-majority dynamics.

One motivation for adopting dynamics in reaching (*simple*) consensus[1] (such as the median dynamics shown in [6]) lies in their provably-good *self-stabilizing* properties against *dynamic adversary corruptions*: it turns out that the 3-majority dynamics has good self-stabilizing properties for the *majority-consensus* problem. More formally, a T-bounded adversary knows the state of every node at the end of each round and, based on this knowledge, he can corrupt the color of up to T nodes in an arbitrary way, just before the next round starts. In this case, the goal is to achieve an almost-stable phase where all but at most $O(T)$ nodes agree on the majority value. This "almost-stability" phase must have poly(n) length,with high probability. Our analysis implicitly shows that the 3-majority dynamics guarantees the self-stabilization property for majority consensus for any k and for $T = o(s/k)$ if the initial bias is $s \geqslant c\sqrt{\min\{2k, (n/\log n)^{1/3}\}\, n \log n}$, for some constant $c > 0$.

[1] In the (simple) consensus problem the goal is to reach any stable monochromatic configuration (any color is accepted) starting from any initial configuration.

Related works. The majority-consensus problem arises in several applications such as distributed database management where data redundancy or replication and majority-rule are used to manage the presence of unknown faulty processors [6, 18]. The objective here is to converge to the version of the data supported by the majority of the initial distributed copies (it is reasonable that a sufficiently large majority of the nodes are not faulty and thus have the correct data). Another application comes from the task of distributed item ranking where every node initially has ranked some item and the goal is to agree on the rank of the item based on the initial majority opinion [19]. Further applications of majority updating-rules in networks can be found in [10, 18].

The results most related to our contribution are those in [6] which have been already discussed above. Several variants of the binary majority consensus have been studied in different distributed models [2, 16]. As for the *population model*, where there is only one random node-pair interaction per round (so the dynamics are strictly sequential), the binary case on the clique has been analyzed in [2] and their generalization to multivalued case ($k \geqslant 3$) does not converge to majority even starting from large bias $s = \Theta(n)$. The polling rule (a somewhat sequential-interaction version of the 1-majority dynamics) has been extensively studied on several classes of graphs (see [18]). More expensive and complex protocols have been considered in order to speed up the process. For instance, in [12], a protocol for the sequential-interaction model is presented that requires $\Theta(\log n)$ memory per nodes and converge in time $O(n^7)$. Other protocols for the sequential-interaction model have been analyzed in [4, 13] (with no time bound). In [1, 3, 7, 19], the polling rule (with 1 more auxiliary state) on the continuous-time population model is proved to converge in $O(n \log n)$ expected time (the bound is not proved in "high probability") (only) when $k = \Theta(1)$ and $s = \Theta(n)$: even assuming such strong restrictions, their analysis, based on real-valued differential-equations, do not work for the discrete-time parallel model considered in this paper. Protocols for specific network topologies and some "social-based" communities have been studied in [1, 7, 15, 19].

Roadmap of the paper. Section 2 formalizes the basic concepts and give some preliminary results. Section 3 is devoted to the proofs of the upper bounds on the converging time of the 3-majority dynamics. In Section 4, the lower bounds for the studied dynamics are described. Section 5 discusses some open questions that we believe to deserve a further study such as the tightness of the initial bias. In the Appendix A, we recall some standard results (such as Chernoff-Bernstein's inequalities, provide a useful probabilistic result on Markov chains (we have not found its explicit proof in the literature) and, finally, we analyze the tightness of our assumptions on the initial bias.

2 Preliminaries

A *k-color distribution*, *k-cd* for short, is any k-tuple $\overline{c} = (c_1, \ldots, c_k)$ such that c_js are non negative integers and $\sum_{j=1,\ldots,k} c_j = n$. A color m is said to be a *majority color* of $\overline{c}$ if $c_m \geqslant c_j$ for every other color $j \in [k] \setminus \{m\}$. We say that $\overline{c}$ is s-biased if a color m exists such that $c_m \geqslant c_j + s$ for every other color $j \in [k] \setminus m$.

The 3-majority protocol runs as follows:

> *At every time step, every node picks three nodes uniformly at random (including itself and with repetitions) and recolors itself according to the majority of the colors it sees. If it sees three different colors, it chooses the first one.*

Clearly, in the case of three different colors, choosing the second or the third one would not make any difference. The same holds even if the choice would be uniformly at random among the three colors.

For any time step t and for any $j \in [k]$, let $C_{j,t}$ be the r. v. counting the number of nodes colored j at time step t and let $C_t = (C_{1,t}, \dots, C_{k,t})$ denote the random variable that is the k-cd at time t of the execution of the 3-majority protocol.

For every $j \in [k]$ let $\mu_j(\overline{c})$ be the expected number of nodes with color j at the next step when the current k-cd is $\overline{c}$,

$$\mu_j(\overline{c}) = \mathbb{E}\left[C_{j,t+1} \mid C_t = \overline{c}\right]$$

Lemma 2.1 (Next expected coloring) *For any k-cd $\overline{c}$ and for every $j \in [k]$, it holds that*

$$\mu_j(\overline{c}) = c_j \left[1 + \frac{1}{n^2}\left(nc_j - \sum_{h\in[k]} c_h^2\right)\right]$$

Proof. According to the 3-majority protocol, a node i gets color j if it chooses three times color j, or if it chooses two times j and one time a different color, or if it chooses the first time color j and then, the second and third time, two different distinct colors. Hence if we name name $X_{i,j}^t$ the indicator random veriable of the event "Node i gets color j at time t", we have that

$$\begin{aligned}
&P\left(X_{i,j}^{t+1} = 1 \mid C_t = \overline{c}\right) \\
&\quad = \left(\frac{c_j}{n}\right)^3 + 3\left(\frac{c_j}{n}\right)^2\left(\frac{n-c_j}{n}\right) + \left(\frac{c_j}{n}\right)\left[1 - \left(\frac{\sum_{h=1}^k c_h^2}{n^2} + 2\left(\frac{c_j}{n}\right)\left(\frac{n-c_j}{n}\right)\right)\right] \\
&\quad = \left(\frac{c_j}{n^3}\right)\left(n^2 + c_j n - \sum_{h=1}^k c_h^2\right)
\end{aligned}$$

□

3 Upper bounds for the 3-majority dynamics

In this section, we show an upper bound on the convergence time of the 3-majority dynamics that holds with high probability. To this aim, we need to evaluate the following random variables. For a k-cd $\overline{c}$, we define

$$\begin{aligned}
m(\overline{c}) &= \max_{h\in[k]} c_h \\
M(\overline{c}) &= \{j \in [k] \mid c_j = m(\overline{c})\} \\
s(\overline{c}) &= \begin{cases} m(\overline{c}) - \max_{h\in[k]-M(\overline{c})} & \text{if } |M(\overline{c})| = 1 \\ 0 & \text{otherwise} \end{cases} \\
\alpha(\overline{c}) &= \frac{(n - m(\overline{c}))s(\overline{c})}{n^2} \\
\gamma(\overline{c}) &= \frac{n \cdot m(\overline{c}) - \sum_{h\in[k]} c_h^2}{n^2} - \alpha(\overline{c})
\end{aligned}$$

The next lemma gives some useful inequalities relating the above quantities.

Lemma 3.1 *For any k-cd $\overline{c}$, the followings hold*

a) $0 \leqslant s(\overline{c}) \leqslant m(\overline{c}) - \frac{n-m(\overline{c})}{k-1}$

b) $0 \leqslant \alpha(\overline{c}) \leqslant \min\{\frac{s(\overline{c})}{n}, \frac{1}{4}\}$

c) $0 \leqslant \gamma(\overline{c}) \leqslant \frac{1}{8}$

Proof. Clearly $s(\overline{c}) \geqslant 0$. If $|M(\overline{c})| > 1$, $s(\overline{c}) = 0$ and the upper bound holds. Otherwise $M(\overline{c}) = \{m\}$ and

$$s(\overline{c}) = m(\overline{c}) - \max\{c_h \mid h \in [k] - \{m\}\} \leqslant c_m - \frac{n - c_m}{k-1} = m(\overline{c}) - \frac{n - m(\overline{c})}{k-1}$$

The only inequality of (b) that is not an immediate consequence of the definition of $\alpha(\overline{c})$ is

$$\alpha(\overline{c}) = \frac{(n - m(\overline{c}))s(\overline{c})}{n^2} \leqslant \frac{(n - m(\overline{c}))m(\overline{c})}{n^2} \leqslant \frac{1}{4}$$

where the last inequality holds since $f(x) = (n-x)x$ has the global maximum at $x = n/2$. As regards (c), let m be any color in $M(\overline{c})$. It holds that

$$\begin{aligned}
\sum_{h\in[k]} c_h^2 &= c_m^2 + \sum_{h\in[k]-\{m\}} c_h^2 \\
&= m(\overline{c})^2 + \sum_{h\in[k]-\{m\}} c_h^2 \\
&\leqslant m(\overline{c})^2 + \left(\sum_{h\in[k]-\{m\}} c_h\right) \max_{h\in[k]-\{m\}} c_h \\
&\leqslant m(\overline{c})^2 + (n - m(\overline{c})) \max_{h\in[k]-\{m\}} c_h \\
&= m(\overline{c})^2 + (n - m(\overline{c}))(m(\overline{c}) - s(\overline{c})) \qquad \text{(since } s(\overline{c}) = m(\overline{c}) - \max_{h\in[k]-\{m\}} c_h\text{)} \\
&= n \cdot m(\overline{c}) - (n - m(\overline{c}))s(\overline{c})
\end{aligned}$$

It immediately follows that $\gamma(\overline{c}) \geqslant 0$. Let $\ell \in [k] - \{m\}$ be such that $c_\ell = \max_{h\in[k]-\{m\}} c_h$. It holds that

$$\begin{aligned}
\gamma(\overline{c}) &= \frac{n \cdot m(\overline{c}) - \sum_{h\in[k]} c_h^2}{n^2} - \alpha(\overline{c}) \\
&= \frac{n \cdot m(\overline{c}) - c_m^2 - c_\ell^2 - \sum_{h\in[k]-\{m,\ell\}} c_h^2 - (n - m(\overline{c}))s(\overline{c})}{n^2} \\
&\leqslant \frac{n \cdot m(\overline{c}) - c_m^2 - c_\ell^2 - (n - m(\overline{c}))s(\overline{c})}{n^2} \\
&= \frac{n \cdot m(\overline{c}) - m(\overline{c})^2 - c_\ell^2 - (n - m(\overline{c}))(m(\overline{c}) - c_\ell)}{n^2} \qquad \text{(since } s(\overline{c}) = m(\overline{c}) - c_\ell\text{)} \\
&= \frac{(n - m(\overline{c}) - c_\ell)c_\ell}{n^2} \\
&= \frac{(n - 2c_\ell - s(\overline{c}))c_\ell}{n^2} \qquad \text{(since } m(\overline{c}) = c_\ell + s(\overline{c})\text{)} \\
&\leqslant \frac{(n - 2c_\ell)c_\ell}{n^2} \\
&\leqslant \frac{1}{8} \qquad \text{(since } f(x) = (n-2x)x \text{ has the global maximum at } x = n/4\text{)}
\end{aligned}$$

□

In the next lemma we provide a new expression for $\mu_j(\overline{c})$ that will be useful in the proofs of Lemmas 3.3 and 3.4.

Lemma 3.2 *Let $\overline{c}$ be any k-cd. Let m be any color in $M(\overline{c})$ and let $\ell \in [k] - \{m\}$ be such that $c_\ell = \max_{h\in[k]-\{m\}} c_h$.*

a) $\mu_m(\overline{c}) = c_m(1 + \gamma(\overline{c}) + \alpha(\overline{c}))$

b) $\forall j \in [k] - M(\overline{c}) \quad \mu_j(\overline{c}) = c_j\left(1 + \gamma(\overline{c}) + \alpha(\overline{c}) - \frac{m(\overline{c})-c_j}{n}\right)$

c) $\mu_\ell(\overline{c}) = c_\ell\left(1 + \gamma(\overline{c}) + \alpha(\overline{c}) - \frac{s(\overline{c})}{n}\right)$

Proof. Equality (a) easily follows from Lemma 2.1 and the definitions of γ and α. As regards equality (b) we have that

$$\begin{aligned}
\mu_j(\overline{c}) &= c_j\left[1 + \frac{1}{n^2}\left(nc_j - \sum_{h\in[k]} c_h^2\right)\right] \qquad \text{(from Lemma 2.1)} \\
&= c_j\left[1 + \frac{1}{n^2}\left(nc_j + n\cdot m(\overline{c}) - n \cdot m(\overline{c}) - \sum_{h\in[k]} c_h^2\right)\right] \\
&= c_j\left[1 + \frac{n \cdot m(\overline{c}) - \sum_{h\in[k]} c_h^2}{n^2} - \alpha(\overline{c}) + \alpha(\overline{c}) - \frac{n\cdot m(\overline{c}) - nc_j}{n^2}\right] \\
&= c_j\left[1 + \gamma(\overline{c}) + \alpha(\overline{c}) - \frac{m(\overline{c}) - c_j}{n}\right]
\end{aligned}$$

Equality (c) follows from (b) by taking into account $s(\overline{c}) = m(\overline{c}) - c_\ell$. □

We now evaluate the increasing rate of the bias of a k-cd during a generic step of the 3-majority dynamics.

Lemma 3.3 (increasing rate of the bias) *Let $\overline{c}$ be any k-cd such that $M(\overline{c}) = \{m\}$ for some $m \in [k]$. Then it holds that, for any $j \in [k] - m$,*

$$\mathbf{P}\left(C_{m,t+1} - C_{j,t+1} \leqslant s(\overline{c}) \left(1 + \gamma(\overline{c}) + \frac{c_m \alpha(\overline{c})}{2s(\overline{c})}\right) \middle| C_t = \overline{c} \right) \leqslant \exp\left(-\frac{c_m \alpha(\overline{c})^2}{25}\right) \quad (1)$$

This is the key-lemma to get our upper bound on the converging time so, before giving its proof, let us provide a rough but useful meaning of Eq. 1 for a fixed setting of parameters k and s, i.e., $k = n^{1/4}$ and $s = c\sqrt{n^{3/4} \log n}$, for some constant $c > 0$. Consider the "initial phase" of the coloring process where c_m is still $\Theta(n/k) = \Theta(n^{3/4})$ and s is still $o(c_m)$. Then, by replacing the values of $\alpha(\overline{c})$ and $\gamma(\overline{c})$ in Eq. 1 (and doing some simple calculations), we get that w.h.p. the bias s increases by a factor $1+\Theta(1/k)$. This is exactly what we need to get the upper bound $O(k \log n)$ on the convergence time. The bound in Eq. 1 has a more complex, general shape since it must work for the whole process and it must lead to our stronger bound $O\left(\min\{k, (n/\log n)^{1/3}\} \log n\right)$.

Proof. (of Lemma 3.3)
In the sequel we tacitly assume that the probabilities, expected values and random variables are conditioned to "$C_t = \overline{c}$". Fix a color $j \in [k] - m$ and let Z be the random variable

$$Z = C_{m,t+1} - C_{j,t+1}$$

It holds that

$$\begin{aligned}
\mathbb{E}[Z] &= \mu_m(\overline{c}) - \mu_j(\overline{c}) \\
&= c_m(1 + \gamma(\overline{c}) + \alpha(\overline{c})) - c_j \left(1 + \gamma(\overline{c}) + \alpha(\overline{c}) - \frac{m(\overline{c}) - c_j}{n}\right) && \text{(from Lemma 3.2)} \\
&= (c_m - c_j)(1 + \gamma(\overline{c})) + c_m \alpha(\overline{c}) + c_j \left(\frac{m(\overline{c}) - c_j}{n} - \alpha(\overline{c})\right) \\
&= (c_m - c_j)(1 + \gamma(\overline{c})) + c_m \alpha(\overline{c}) + c_j \left(\frac{m(\overline{c}) - c_j}{n} - \alpha(\overline{c})\right) \\
&\geqslant (c_m - c_j)(1 + \gamma(\overline{c})) + c_m \alpha(\overline{c}) + c_j \left(\frac{s(\overline{c})}{n} - \alpha(\overline{c})\right) && \text{(since } m(\overline{c}) - c_j \geqslant s(\overline{c})\text{)} \\
&\geqslant (c_m - c_j)(1 + \gamma(\overline{c})) + c_m \alpha(\overline{c}) && \text{(since } \tfrac{s(\overline{c})}{n} \geqslant \alpha(\overline{c}) \text{ by Lemma 3.1)} \\
&\geqslant s(\overline{c}) \left(1 + \gamma(\overline{c}) + \frac{c_m \alpha(\overline{c})}{s(\overline{c})}\right) && \text{(since } c_m - c_j \geqslant s(\overline{c})\text{)} \qquad (2)
\end{aligned}$$

In order to make use of the Bernstein's Inequality (see Lemma A.4 in Appendix A.1) we introduce, for any $i \in [n]$, the random variable

$$Z_i = \begin{cases} 1 & \text{if node } i \text{ gets color } m \text{ at time } t+1 \\ -1 & \text{if node } i \text{ gets color } j \text{ at time } t+1 \\ 0 & \text{otherwise} \end{cases}$$

Clearly, the Z_i's are independent and it holds that

$$Z = \sum_{i \in [n]} Z_i$$

In order to apply the Bernstein's Inequality (Lemma A.4) to $-Z$, we firstly observe that

$$-Z_i - \mathbb{E}[-Z_i] \leqslant 2$$

so we can choose $b = 2$. As for the variance σ^2 of $-Z$, we have that

$$\begin{aligned}
\sigma^2 = \mathbf{Var}[-Z] &= \sum_{i\in[n]} \mathbf{Var}[-Z_i] \\
&= \sum_{i\in[n]} \left(\mathbb{E}\left[(-Z_i)^2\right] - \mathbb{E}[-Z_i]^2\right) = \sum_{i\in[n]} \left(\mathbb{E}\left[Z_i^2\right] - \mathbb{E}[Z_i]^2\right) \\
&\leqslant \sum_{i\in[n]} \mathbb{E}\left[Z_i^2\right] = \sum_{i\in[n]} \left(\mathbf{P}(Z_i = 1) + \mathbf{P}(Z_i = -1)\right) = \mu_m(\overline{c}) + \mu_j(\overline{c}) \\
&\leqslant 2\mu_m(\overline{c}) \qquad \text{(since } \mu_j(\overline{c}) \leqslant \mu_m(\overline{c}) \text{ by Lemma 3.2)} \\
&= 2c_m(1 + \gamma(\overline{c}) + \alpha(\overline{c})) \\
&\leqslant 2c_m\left(1 + \frac{1}{8} + \frac{1}{4}\right) \qquad \text{(from Lemma 3.1)} \\
&\leqslant 3c_m
\end{aligned} \tag{3}$$

For the sake of convenience, let us define

$$P = \mathbf{P}\left(Z < s(\overline{c})\left(1 + \gamma(\overline{c}) + \frac{c_m\alpha(\overline{c})}{2s(\overline{c})}\right)\right)$$

Now we conclude the proof by applying the Bernstein's Inequality

$$\begin{aligned}
P &= \mathbf{P}\left(-Z > -s(\overline{c})\left(1 + \gamma(\overline{c}) + \frac{c_m\alpha(\overline{c})}{s(\overline{c})}\right) + \frac{c_m\alpha(\overline{c})}{2}\right) \\
&\leqslant \mathbf{P}\left(-Z > \mathbb{E}[-Z] + \frac{c_m\alpha(\overline{c})}{2}\right) \qquad \left(\mathbb{E}[-Z] \leqslant -s(\overline{c})\left(1 + \gamma(\overline{c}) + \tfrac{c_m\alpha(\overline{c})}{s(\overline{c})}\right) \text{ by Ineq. 2}\right) \\
&\leqslant \exp\left(-\frac{\left(\frac{c_m\alpha(\overline{c})}{2}\right)^2}{2\sigma^2 + (4/3)\frac{c_m\alpha(\overline{c})}{2}}\right) \qquad \left(\text{from Bernstein's Inequalty, Lemma A.4 with } b = 2 \text{ and } \lambda = \tfrac{c_m\alpha(\overline{c})}{2}\right) \\
&\leqslant \exp\left(-\frac{c_m^2\alpha(\overline{c})^2}{24c_m + (8/3)c_m\alpha(\overline{c})}\right) \qquad \text{(from Ineq. 3)} \\
&\leqslant \exp\left(-\frac{c_m\alpha(\overline{c})^2}{25}\right) \qquad \text{(since } \alpha(\overline{c}) \leqslant 1/4 \text{ by Lemma 3.1)}
\end{aligned}$$

□

The next lemma derives from Lemmas 3.1 and 3.2.

Lemma 3.4 *Let $\overline{c}$ be any k-cd such that $M(\overline{c}) = \{m\}$ for some $m \in [k]$. It holds that*

$$\mathbf{P}\left(C_{m,t+1} \leqslant c_m\left(1 + \gamma(\overline{c}) + \frac{\alpha(\overline{c})}{2}\right) \,\middle|\, C_t = \overline{c}\right) \leqslant \exp\left(-\frac{c_m\alpha(\overline{c})^2}{11}\right)$$

Proof. Let

$$P_m = \mathbf{P}\left(C_{m,t+1} \leqslant c_m\left(1 + \gamma(\overline{c}) + \frac{\alpha(\overline{c})}{2}\right) \,\middle|\, C_t = \overline{c}\right)$$

and let

$$\delta_m = \frac{\alpha(\overline{c})}{2(1+\gamma(\overline{c})+\alpha(\overline{c}))}$$

From Lemma 3.1, $\gamma, \alpha \geqslant 0$, and thus $0 < \delta_m < 1$. Thanks to the Chernoff bound we have that

$$\begin{aligned}
P_m &= \mathbf{P}\left(C_{m,t+1} \leqslant (1-\delta_m)\mu_m \,|\, C_t = \overline{c}\right) \quad \text{(from Lemma 3.2)} \\
&\leqslant \exp\left(-\frac{\delta_m^2 \mu_m}{2}\right) \qquad \text{(by the Chernoff bound)} \\
&= \exp\left(-\frac{1}{2}\left(\frac{\alpha(\overline{c})}{2(1+\gamma(\overline{c})+\alpha(\overline{c}))}\right)^2 c_m(1+\gamma(\overline{c})+\alpha(\overline{c}))\right) \\
&= \exp\left(-\frac{c_m \alpha(\overline{c})^2}{8(1+\gamma(\overline{c})+\alpha(\overline{c}))}\right) \\
&\leqslant \exp\left(-\frac{c_m \alpha(\overline{c})^2}{11}\right) \qquad \text{(since } \gamma(\overline{c})+\alpha(\overline{c}) \leqslant 3/8 \text{ by Lemma 3.1)}
\end{aligned}$$

□

We now use Lemmas 3.3 and 3.4 in order to get some bounds on the increasing rate of the bias: they will lead to a bound on convergence time that does not depend on k.

Lemma 3.5 (large majority and large bias) *Let $\overline{c}$ be any k-cd such that $M(\overline{c}) = \{m\}$ for some $m \in [k]$. For any value λ with $0 < \lambda \leqslant 2/3$, if $\lambda n \leqslant c_m \leqslant (2/3)n$ and $s(\overline{c}) \geqslant 22\sqrt{(1/\lambda)n \log n}$, then, for every $j \in [k] - \{m\}$,*

$$\mathbf{P}\left(C_{m,t+1} - C_{j,t+1} \leqslant s(\overline{c})\left(1+\frac{\lambda}{6}\right) \,\middle|\, C_t = \overline{c}\right) \;\leqslant\; \frac{1}{n^2}$$

and

$$\mathbf{P}\left(C_{m,t+1} \leqslant c_m \,|\, C_t = \overline{c}\right) \;\leqslant\; \frac{1}{n^2}$$

Proof. From Lemma 3.3 we have that

$$\mathbf{P}\left(C_{m,t+1} - C_{j,t+1} \leqslant s(\overline{c})\left(1+\gamma(\overline{c})+\frac{c_m\alpha(\overline{c})}{2s(\overline{c})}\right) \,\middle|\, C_t = \overline{c}\right) \;\leqslant\; \exp\left(-\frac{c_m\alpha(\overline{c})^2}{25}\right) \tag{4}$$

It holds that

$$\frac{c_m\alpha(\overline{c})}{2s(\overline{c})} = \frac{c_m(n-c_m)}{2n^2} \geqslant \frac{\lambda n(n/3)}{2n^2} \geqslant \frac{\lambda}{6} \tag{5}$$

As regards the exponent of the probability bound of Ineq. 4 we get

$$\begin{aligned}
\frac{c_m\alpha(\overline{c})^2}{25} &= \frac{c_m(n-c_m)^2 s(\overline{c})^2}{25n^4} \\
&\geqslant \frac{c_m s(\overline{c})^2}{225n^2} \qquad \text{(since } c_m \leqslant (2/3)n\text{)} \\
&\geqslant \frac{\lambda n 484(1/\lambda)n\log n}{225n^2} \qquad \text{(by the hypothesis bounds on } c_m \text{ and } s(\overline{c})\text{)} \\
&\geqslant 2\log n
\end{aligned} \tag{6}$$

By combining Ineq.s 4, 5, and 6 we obtain the first probability bound. As for the second bound, from Lemma 3.4 it holds that

$$\mathbf{P}\left(C_{m,t+1} \leqslant c_m \mid C_t = \overline{c}\right) \leqslant \mathbf{P}\left(C_{m,t+1} \leqslant c_m \left(1 + \gamma(\overline{c}) + \frac{\alpha(\overline{c})}{2}\right) \,\middle|\, C_t = \overline{c}\right) \leqslant \exp\left(-\frac{c_m \alpha(\overline{c})^2}{11}\right)$$

and, from Ineq. 6,

$$\frac{c_m \alpha(\overline{c})^2}{11} \geqslant \frac{c_m \alpha(\overline{c})^2}{25} \geqslant 2 \log n$$

□

For any $m \in [k]$, let $\overline{C}_{m,t} = n - C_{m,t}$ denote the random variable counting the number of nodes with colors different from m at time t. For any k-cd $\overline{c}$ and for any $m \in [k]$, let

$$\overline{\mu}_m(\overline{c}) = \mathbb{E}\left[\overline{C}_{m,t+1} \,\middle|\, C_t = \overline{c}\right]$$

Lemma 3.6 *For any k-cd $\overline{c}$ and for any $m \in M(\overline{c})$, it holds that*

$$(n - c_m)\left(1 - \frac{c_m^2}{n^2}\right) \leqslant \overline{\mu}_m(\overline{c}) \leqslant (n - c_m)\left(1 - \frac{s(\overline{c}) c_m}{n^2}\right)$$

Proof. Firstly we observe that

$$\begin{aligned}
\overline{\mu}_m(\overline{c}) &= n - \mu_m(\overline{c}) \\
&= n - c_m(1 + \gamma(\overline{c}) + \alpha(\overline{c})) \qquad \text{(from Lemma 3.2)} \\
&= n - c_m - c_m \frac{n c_m - \sum_{h \in [k]} c_h^2}{n^2} \\
&= n - c_m - c_m \frac{n c_m - c_m^2 - \sum_{h \in [k]-\{m\}} c_h^2}{n^2} \\
&= n - c_m - \frac{c_m^2}{n^2}(n - c_m) + \frac{c_m \sum_{h \in [k]-\{m\}} c_h^2}{n^2} \\
&= (n - c_m)\left(1 - \frac{c_m^2}{n^2}\right) + \frac{c_m \sum_{h \in [k]-\{m\}} c_h^2}{n^2} \\
&= \frac{(n - c_m)^2 (n + c_m)}{n^2} + \frac{c_m \sum_{h \in [k]-\{m\}} c_h^2}{n^2}
\end{aligned} \tag{7}$$

Then, by taking into account the following bound

$$\sum_{h \in [k]-\{m\}} c_h^2 \leqslant \sum_{h \in [k]-\{m\}} c_h \max_{j \in [k]-\{m\}} c_j = (n - c_m) \max_{j \in [k]-\{m\}} c_j \leqslant (n - c_m)(c_m - s(\overline{c}))$$

we obtain

$$\overline{\mu}_m(\overline{c}) \leqslant \frac{(n - c_m)^2 (n + c_m)}{n^2} + \frac{c_m (n - c_m)(c_m - s(\overline{c}))}{n^2} = (n - c_m)\left(1 - \frac{s(\overline{c}) c_m}{n^2}\right)$$

This proves the right-hand inequality. The left-hand inequality is immediate from Eq. 7

$$\overline{\mu}_m(\overline{c}) = \frac{(n - c_m)^2 (n + c_m)}{n^2} + \frac{c_m \sum_{h \in [k]-\{m\}} c_h^2}{n^2} \geqslant \frac{(n - c_m)^2 (n + c_m)}{n^2} = (n - c_m)\left(1 - \frac{c_m^2}{n^2}\right)$$

□

In the next lemma we show that, when the bias of a k-cd is at least $n/3$, then the number of nodes that do not have the majority color decreases at an exponential rate.

Lemma 3.7 (very-large majority) *Let $\overline{c}$ be any k-cd such that $s(\overline{c}) \geqslant n/3$ for some $m \in [k]$. If $n - c_m \geqslant \sqrt[4]{n}\log n$, then*

$$\mathbf{P}\left(\overline{C}_{m,t+1} \geqslant \frac{17}{18}(n-c_m) \,\middle|\, C_t = \overline{c}\right) \leqslant \frac{1}{n^2}$$

If $n - c_m < \sqrt[4]{n}\log n$, then

$$\mathbf{P}\left(\overline{C}_{m,t+1} > 0 \,\middle|\, C_t = \overline{c}\right) \leqslant \frac{1}{\sqrt[5]{n}} \qquad \text{and} \qquad \mathbf{P}\left(\overline{C}_{m,t+1} \geqslant \sqrt[4]{n}\log n \,\middle|\, C_t = \overline{c}\right) \leqslant \frac{1}{n^2}$$

Proof. For any $i \in [n]$, let X_i be the random variable defined as

$$X_i = \begin{cases} 1 & \text{if node } i \text{ does not get colour } m \text{ at time } t+1 \text{ under the condition } C_t = \overline{c} \\ 0 & \text{otherwise} \end{cases}$$

In the sequel of the proof we tacitly assume the condition $C_t = \overline{c}$. Clearly the X_is are independent and

$$\overline{C}_{m,t+1} = \sum_{i\in[n]} X_i$$

Notice that $s(\overline{c}) \geqslant n/3$ implies $c_m \geqslant n/3$. For the case $n - c_m \geqslant \sqrt[4]{n}\log n$, we distinguish two sub-cases. Firstly assume that $n - c_m \geqslant 34\sqrt{n\log n}$. Observe that

$$\begin{aligned}
\overline{\mu}_m(\overline{c}) &\leqslant (n-c_m)\left(1 - \frac{s(\overline{c})c_m}{n^2}\right) && \text{(from Lemma 3.6)} \\
&= (n-c_m)\left(1 - \frac{\frac{n}{3}\frac{n}{3}}{n^2}\right) && \text{(since } s(\overline{c}) \geqslant n/3) \\
&\leqslant \frac{8}{9}(n-c_m) && (8)
\end{aligned}$$

Let $\delta = 1/16$, it holds that

$$\begin{aligned}
\mathbf{P}\left(\overline{C}_{m,t+1} \geqslant \frac{17}{18}(n-c_m)\right) &\leqslant \mathbf{P}\left(\overline{C}_{m,t+1} \geqslant (1+\delta)\overline{\mu}_m(\overline{c})\right) && \text{(from Ineq. 8)} \\
&\leqslant \exp\left(-\frac{\delta^2\overline{\mu}_m(\overline{c})}{3}\right) && \text{(from the Chernoff Bound)} \\
&\leqslant \exp\left(-\frac{\delta^2}{3}(n-c_m)\left(1-\frac{c_m^2}{n^2}\right)\right) && \text{(from Lemma 3.6)} \\
&= \exp\left(-\frac{\delta^2(n-c_m)^2(n+c_m)}{3n^2}\right) \\
&\leqslant \exp\left(-\frac{\delta^2 1156 n\log n(4/3)n}{3n^2}\right) \\
&\qquad \text{(since } n-c_m \geqslant 34\sqrt{n\log n} \text{ and } c_m \geqslant n/3) \\
&\leqslant \frac{1}{n^2}
\end{aligned}$$

Now, consider the sub-case $n - c_m < 34\sqrt{n\log n}$ and of course $n - c_m \geqslant \sqrt[4]{n}\log n$. In this sub-case it holds that

$$c_m > n - 34\sqrt{n\log n} \qquad \text{and} \qquad s(\overline{c}) > n - 68\sqrt{n\log n}$$

Thus, we have that

$$\begin{aligned}
\overline{\mu}_m(\overline{c}) &\leqslant (n-c_m)\left(1-\frac{s(\overline{c})c_m}{n^2}\right) \qquad \text{(from Lemma 3.6)}\\
&\leqslant (n-c_m)\left(1-\frac{(n-68\sqrt{n\log n})(n-34\sqrt{n\log n})}{n^2}\right)\\
&\leqslant 102\sqrt{\frac{\log n}{n}}(n-c_m) \qquad (9)
\end{aligned}$$

Let

$$\delta = \frac{1}{108}\sqrt{\frac{n}{\log n}} - 1$$

It holds that

$$\begin{aligned}
\mathbf{P}\left(\overline{C}_{m,t+1} \geqslant \frac{17}{18}(n-c_m)\right) &\leqslant \mathbf{P}\left(\overline{C}_{m,t+1} \geqslant (1+\delta)\overline{\mu}_m(\overline{c})\right) \qquad \text{(from Ineq. 9)}\\
&\leqslant \exp\left(-\delta\overline{\mu}_m(\overline{c})\right) \qquad \text{(from the Chernoff Bound)}\\
&\leqslant \exp\left(-\left(\frac{1}{108}\sqrt{\frac{n}{\log n}}-1\right)\overline{\mu}_m(\overline{c})\right)\\
&\leqslant \exp\left(-\frac{1}{200}\sqrt{\frac{n}{\log n}}\overline{\mu}_m(\overline{c})\right) \qquad \text{(for sufficiently large } n\text{)}\\
&\leqslant \exp\left(-\frac{1}{200}\sqrt{\frac{n}{\log n}}(n-c_m)\left(1-\frac{c_m^2}{n^2}\right)\right) \qquad \text{(from Lemma 3.6)}\\
&= \exp\left(-\frac{1}{200}\sqrt{\frac{n}{\log n}}\frac{(n-c_m)^2(n+c_m)}{n^2}\right)\\
&\leqslant \exp\left(-\frac{1}{200}\sqrt{\frac{n}{\log n}}\sqrt{n}\frac{\log^2 n}{n}\right) \qquad \text{(since } n-c_m \geqslant \sqrt[4]{n}\log n\text{)}\\
&\leqslant \exp\left(-\frac{1}{200}\log n\sqrt{\log n}\right)\\
&\leqslant \frac{1}{n^2} \qquad \text{(for sufficiently large } n\text{)}
\end{aligned}$$

Consider now the case $n - c_m < \sqrt[4]{n}\log n$. We assume that $c_m > 0$ otherwise the probability bounds are trivially true. It holds that

$$\begin{aligned}
\overline{\mu}_m(\overline{c}) &\leqslant (n-c_m)\left(1-\frac{s(\overline{c})c_m}{n^2}\right) \qquad \text{(from Lemma 3.6)}\\
&\leqslant 3\frac{\log^2 n}{\sqrt{n}} \qquad \text{(since } n-c_m \leqslant \sqrt[4]{n}\log n\text{)} \qquad (10)
\end{aligned}$$

Let $\delta = 1/(2\overline{\mu}_m(\overline{c})) - 1$. We have that

$$\begin{aligned}
\mathbf{P}\left(\overline{C}_{m,t+1} > 0\right) &= \mathbf{P}\left(\overline{C}_{m,t+1} > (1+\delta)\overline{\mu}_m(\overline{c})\right) \qquad \text{(since } \overline{C}_{m,t+1} \text{ has integer values only)} \\
&\leqslant \left(\frac{e^{\delta}}{(1+\delta)^{1+\delta}}\right)^{\overline{\mu}_m(\overline{c})} \qquad \text{(from the Chernoff bound)} \\
&= \left(\frac{e^{\frac{1}{2\overline{\mu}_m(\overline{c})}-1}}{\left(\frac{1}{2\overline{\mu}_m(\overline{c})}\right)^{\frac{1}{2\overline{\mu}_m(\overline{c})}}}\right)^{\overline{\mu}_m(\overline{c})} \\
&= \sqrt{2e\overline{\mu}_m(\overline{c})}e^{-\overline{\mu}_m(\overline{c})} \\
&\leqslant \sqrt{6e\frac{\log^2 n}{\sqrt{n}}} \qquad \text{(from Ineq. 10)} \\
&\leqslant \frac{1}{\sqrt[5]{n}} \qquad \text{(for sufficiently large } n\text{)}
\end{aligned}$$

As regards the last bound, let

$$\delta = \frac{\sqrt[4]{n}\log n}{2\overline{\mu}_m(\overline{c})} - 1$$

It holds that

$$\begin{aligned}
\mathbf{P}\left(\overline{C}_{m,t+1} > \sqrt[4]{n}\log n\right) &\leqslant \mathbf{P}\left(\overline{C}_{m,t+1} > (1+\delta)\overline{\mu}_m(\overline{c})\right) \\
&\leqslant \exp(-\delta\overline{\mu}_m(\overline{c})) \qquad \text{(from the Chernoff bound)} \\
&= \exp\left(-\frac{\sqrt[4]{n}\log n}{2} + \overline{\mu}_m(\overline{c})\right) \\
&\leqslant \exp\left(-\frac{\sqrt[4]{n}\log n}{2} + 3\frac{\log^2 n}{\sqrt{n}}\right) \qquad \text{(from Ineq. 10)} \\
&\leqslant \frac{1}{n^2} \qquad \text{(for sufficiently large } n\text{)}
\end{aligned}$$

□

The main result of this section can be proved by using Lemmas 3.5 and 3.7.

Theorem 3.8 (the general upper bound) *Let λ be any value such that $3 \leqslant \lambda < \sqrt{n}$. If $\overline{c}$ is a k-cd such that, for some $m \in [k]$, $M(\overline{c}) = \{m\}$, $c_m \geqslant n/\lambda$, and $s(\overline{c}) \geqslant 22\sqrt{\lambda\, n \log n}$, then w.h.p. the 3-majority protocol converges to color m in $\mathrm{O}(\lambda \log n)$ time.*

Proof. For the sake of convenience, let

$$\Lambda = 22\sqrt{\lambda\, n \log n}$$

Notice that $\Lambda \leqslant 22n^{3/4}\sqrt{\log n}$. In order to make use of Lemma A.5 (see the appendix), we consider the Markov chain determined by the 3-majority protocol. The states of the Markov chain are all the possible assignments of the k colors to the n nodes. For any assignment $\mathbf{a}$, let $\overline{c}(\mathbf{a})$ denote the k-cd determined by $\mathbf{a}$ and let $c_j(\mathbf{a})$ denote any its component. Let X_t be the random variable that is the state at time t given that X_0 is a state whose k-cd is $\overline{c}$. Define

$$T_1 = \left\lfloor 1 + \frac{\log \frac{n}{3\Lambda}}{\log\left(1+\frac{1}{6\,\lambda}\right)} \right\rfloor$$
$$T_2 = \left\lfloor 1 + \frac{3}{2\log(18/17)} \log \frac{n^{3/4}}{\log n} \right\rfloor$$

For any $i = 1, \dots, T_1$, let

$$A_i = \left\{ \mathbf{a} \;\middle|\; c_m(\mathbf{a}) > \frac{2}{3} n \vee \left(M(\overline{c}(\mathbf{a})) = \{m\} \wedge c_m(\mathbf{a}) \geqslant \frac{n}{\lambda} \wedge s(\overline{c}(\mathbf{a})) \geqslant \Lambda \left(1 + \frac{1}{6\,\lambda}\right)^{i-1} \right) \right\}$$

Observe that $X_0 \in A_1$ and $A_1 \supseteq A_2 \supseteq \cdots \supseteq A_{T_1}$. For any $i = 1, \dots, T_2$, let

$$A_{T_1+i} = \left\{ \mathbf{a} \;\middle|\; M(\overline{c}(\mathbf{a})) = \{m\} \wedge s(\overline{c}(\mathbf{a})) \geqslant \frac{n}{3} \wedge n - c_m(\mathbf{a}) \leqslant \frac{2n}{3} \left(\frac{17}{18}\right)^{i-1} \right\}$$

and let

$$A_{T_1+T_2+1} = \left\{ \mathbf{a} \;\middle|\; M(\overline{c}(\mathbf{a})) = \{m\} \wedge s(\overline{c}(\mathbf{a})) \geqslant \frac{n}{3} \wedge n - c_m(\mathbf{a}) < \sqrt[4]{n} \log n \right\}$$
$$A_{T_1+T_2+2} = \{\mathbf{a} \mid c_m(\mathbf{a}) = n\}$$

It is easy to verify that $A_{T_1} \supseteq A_{T_1+1} \supseteq A_{T_1+2} \supseteq \cdots \supseteq A_{T_1+T_2} \supseteq A_{T_1+T_2+1} \supseteq A_{T_1+T_2+2}$. Thus it holds that $A_1 \supseteq \cdots \supseteq A_{T_1+T_2+2}$. Taking into account that $c_m(\mathbf{a}) > (2/3)n$ implies $s(\overline{c}(\mathbf{a})) \geqslant n/3$, from Lemma 3.5 we have that, for any $i = 1, \dots, T_1$,

$$\mathbf{P}\left(X_t \in A_i \mid X_{t-1} \in A_i\right) \geqslant 1 - \frac{1}{n^2} \quad \text{and} \quad \mathbf{P}\left(X_t \in A_{i+1} \mid X_{t-1} \in A_i\right) \geqslant 1 - \frac{1}{n^2}$$

From Lemma 3.7 we get, for any $i = T_1 + 1, \dots, T_1 + T_2$

$$\mathbf{P}\left(X_t \in A_i \mid X_{t-1} \in A_i\right) \geqslant 1 - \frac{1}{n^2} \quad \text{and} \quad \mathbf{P}\left(X_t \in A_{i+1} \mid X_{t-1} \in A_i\right) \geqslant 1 - \frac{1}{n^2}$$

moreover

$$\mathbf{P}\left(X_t \in A_{T_1+T_2+1} \mid X_{t-1} \in A_{T_1+T_2+1}\right) \geqslant 1 - \frac{1}{n^2} \quad \text{and}$$
$$\mathbf{P}\left(X_t \in A_{T_1+T_2+2} \mid X_{t-1} \in A_{T_1+T_2+1}\right) \geqslant 1 - \frac{1}{\sqrt[5]{n}}$$

Hence, by applying Lemma A.5 with $\epsilon = 1/n^2$ and $\nu = 1/\sqrt[5]{n}$ with $\ell = 10$, we obtain

$$\mathbf{P}\left(X_{10T} \in A_T \mid X_0 \in A_1\right) \geqslant 1 - T\left(\frac{10}{n^2} + \left(\frac{1}{\sqrt[5]{n}}\right)^{10}\right) = 1 - \frac{11T}{n^2}$$

It easy to see that $T < n/11$. Thus w.h.p. in time $10T$ the 3-majority protocol converges to color m. Now we bound more precisely T. It holds that

$$
\begin{aligned}
T &= T_1 + T_2 + 2 \\
&\leqslant 4 + \frac{\log \frac{n}{3\Lambda}}{\log\left(1+\frac{1}{6\lambda}\right)} + \frac{3}{2\log(18/17)} \log \frac{n^{3/4}}{\log n} \\
&\leqslant 4 + 27\log n + \frac{\log \frac{n}{66\sqrt{\lambda\, n \log n}}}{\log\left(1+\frac{1}{6\lambda}\right)} \\
&\leqslant 28\log n + \frac{\log\left(\frac{1}{66}\sqrt{\frac{n}{\lambda \log n}}\right)}{\frac{1/(6\lambda)}{1+1/(6\lambda)}} \qquad \left(\text{since } \log(1+x) \geqslant \tfrac{x}{1+x}\right) \\
&\leqslant 26\log n + 7\lambda \log(n/\lambda) \\
&\leqslant 10\,\lambda\,\log n
\end{aligned}
$$

□

Observation 3.9 *Let us consider a dynamic adversary (see the Introduction) that can change the color of up to T nodes at the beginning of each time step and assume $T = o(\lambda \cdot s)$. Then, Theorem 3.8 still holds since the impact of such a T-bounded adversary is negligible in the growth of the bias s (this can be easily seen in the proof of Lemma 3.5). For instance, when $k \leqslant 2\sqrt[3]{\frac{n}{\log n}}$, then the tolerance of the 3-majority dynamics is $T = o(s/k)$.*

The next three corollaries of Theorem 3.8 address three relevant special cases. Corollary 3.10 is obtained by setting $\lambda = \min\left\{2k,\ \sqrt[3]{\frac{n}{\log n}}\right\}$ and it provides a bound which does not assume any condition on c_m.

Corollary 3.10 *If $\overline{c}$ is a k-cd such that, for some $m \in [k]$, $M(\overline{c}) = \{m\}$ and*

$$
s(\overline{c}) \geqslant 22\sqrt{\min\left\{2k,\ \sqrt[3]{\frac{n}{\log n}}\right\} n \log n}
$$

then, w.h.p. the 3-majority protocol converges to color m in $\mathrm{O}\left(\min\left\{2k,\ \sqrt[3]{\frac{n}{\log n}}\right\}\log n\right)$ time.

Corollaries 3.11 and 3.12 are obtained by setting $\lambda = \operatorname{poly}\log(n)$ and $\lambda = \Theta(1)$, respectively. They require some lower bounds on c_m.

Corollary 3.11 *If $\overline{c}$ is a k-cd such that, for some $m \in [k]$, $M(\overline{c}) = \{m\}$, $c_m \geqslant n/\log^{\ell} n$, and $s(\overline{c}) \geqslant 22\sqrt{n \log^{\ell+1} n}$, then w.h.p. the 3-majority protocol converges to color m in $\mathrm{O}(\log^{\ell+1} n)$ time.*

Corollary 3.12 *If $\overline{c}$ is a k-cd such that, for some $m \in [k]$, $M(\overline{c}) = \{m\}$, $c_m \geqslant n/\beta$, and $s(\overline{c}) \geqslant 22\sqrt{\beta n \log n}$, for some constant $\beta \geqslant 3$, then w.h.p. the 3-majority protocol converges to color m in $\mathrm{O}(\log n)$ time.*

4 The lower bounds

This section is organized in 3 subsections: in the first one, we prove a lower bound on the convergence time of the 3-majority dynamics; in the second subsection, we show that any 3-input dynamics, that does not work as 3-majority, cannot converge to majority consensus; finally, in the third subsection, we provide a lower bound on the convergence time of the general h-majority.

4.1 A lower bound for the 3-majority dynamics

In this section we show that if the 3-majority dynamics starts from a sufficiently balanced configuration (i.e., at the beginning there are $n/k \pm o(n/k)$ nodes of every color) then it will take $\Omega(k \log n)$ steps w.h.p. to reach one of the absorbing configurations where all nodes have the same color. In what follows, all events and random variables thus concern the Markovian process yielded by the 3-majority dynamics.

In the next lemma we show that if there are at most $n/k + b$ nodes of a specific color, where b is smaller than n/k, then at the next time step there are at most $n/k + (1 + 3/k)b$ nodes of that color w.h.p.

Lemma 4.1 *Let the number of colors k be such that $k \leqslant (n/\log n)^{1/4}$, let b be any number with $k\sqrt{n \log n} \leqslant b \leqslant n/k$, and let $\{X_t\}$ be the sequence of random variables where X_t is the number of a specific color at time t. If $X_t = n/k+a$ for some $a \leqslant b$ then $X_{t+1} \leqslant n/k+(1+3/k)b$ w.h.p.; more precisely, for any $a \leqslant b$ it holds that*

$$\mathbf{P}\left(X_{t+1} \geqslant \frac{n}{k} + \left(1 + \frac{3}{k}\right) b \;\middle|\; X_t = \frac{n}{k} + a\right) \leqslant \frac{1}{n^2}$$

Proof. For a color h and time step t, let $C_{h,t}$ be the random variable indicating the number of nodes with color h, let $C_t = (C_{1,t}, \ldots, C_{k,t})$ be the random variable indicating the *coloring* at time t. For any coloring $\overline{c} = (c_1, \ldots, c_k)$ with $\sum_{h=1}^k c_h = n$ and any color $h \in [k]$, the expected value of the number of nodes colored h at time $t+1$ given $C_t = \overline{c}$ is (see Lemma 2.1)

$$\mathbf{E}\left[C_{h,t+1} \mid C_t = \overline{c}\right] = c_{h,t}\left(1 + \frac{c_{h,t}}{n} - \frac{1}{n^2}\sum_{j=1}^{k} c_j^2\right)$$

Observe that, since $\sum_{j=1}^k c_j = n$, from Jensen inequality (see Lemma A.3) it follows that $(1/n^2)\sum_{j=1}^k c_j^2 \geqslant 1/k$. Hence, if X_t is the random variable counting the number of nodes of one specific color, then we can give an upper bound on the expectation of X_{t+1} that depends only on X_t and not on the whole coloring at time t, namely

$$\mathbf{E}\left[X_{t+1} \mid X_t\right] \leqslant X_t\left(1 + \frac{X_t}{n} - \frac{1}{k}\right)$$

If we condition on the number of nodes of that specific color being of the form $n/k + a$ for some $a \leqslant b$ we get

$$\begin{aligned}
\mathbf{E}\left[X_{t+1} \,|\, X_t = n/k + a\right] &\leqslant \left(\frac{n}{k} + a\right)\left(1 + \frac{n/k + a}{n} - \frac{1}{k}\right) \\
&= \left(\frac{n}{k} + a\right)\left(1 + \frac{a}{n}\right) = \frac{n}{k} + \left(1 + \frac{1}{k}\right) a + \frac{a^2}{n} \\
&\leqslant \frac{n}{k} + \left(1 + \frac{1}{k}\right) b + \frac{b^2}{n} \leqslant \frac{n}{k} + \left(1 + \frac{2}{k}\right) b
\end{aligned}$$

where in the last two inequalities we used that $a \leqslant b$ and $b \leqslant n/k$.[2] Since X_t can be written as a sum of n independent Bernoulli random variables, from Chernoff bound (see Lemma A.2) we thus get that for every $a \leqslant b$ it holds that

$$\mathbf{P}\left(X_{t+1} \geqslant \frac{n}{k} + \left(1 + \frac{3}{k}\right) b \,|\, X_t = n/k + a\right) \leqslant e^{-2(b/k)^2/n} \leqslant \frac{1}{n^2}$$

where in the last inequality we used that $b \geqslant k\sqrt{n \log n}$. $\square$

Let us say that a coloring $\overline{c} = (c_1, \ldots, c_k) \in \{0, 1, \ldots, n\}^k$ with $\sum_{h=1}^{k} c_h = n$ is *monochromatic* if there is an $h \in [k]$ such that $c_h = n$. In the next theorem we show that if we start from a sufficiently *balanced* coloring, then the 3-majority protocol takes $\Omega(k \log n)$ time steps w.h.p. to reach a monochromatic coloring.

Theorem 4.2 *Let C_t be the random variable indicating the coloring at time t according to the 3-majority protocol and let $\tau = \inf\{t \in \mathbb{N} : C_t \text{ is monochromatic}\}$ be the random variable indicating the first time step such that C_t is monochromatic. If the initial number of colors is $k \leqslant (n/\log n)^{1/4}$ and the initial coloring is $C_0 = (c_1, \ldots, c_k)$ with $\max\{c_h : h = 1, \ldots, k\} \leqslant \frac{n}{k} + \left(\frac{n}{k}\right)^{1-\varepsilon}$ then $\tau = \Omega(k \log n)$ w.h.p.*

> *Idea of the proof.* For a color $h \in [k]$ let us name *positive unbalance* the difference $C_{h,t} - n/k$ between the number of nodes colored h and the ratio of the total number of nodes and the total number of colors. In Lemma 4.1 we proved that as long as the positive unbalance of a color is smaller than n/k, it will increase by a factor smaller than $(1 + 3/k)$ w.h.p. at every time step. Hence, if a color starts with a positive unbalance smaller than $(n/k)^{1-\varepsilon}$ it will take $\Omega(k \log n)$ time steps to reach an unbalance of n/k w.h.p. By union bounding on all the colors we get the result. A full-detailed proof follows. $\square$

Proof. Observe that if $T \leqslant c\,k \log n$, for a suitable positive constant c, then $(1-3/k)^T (n/k)^{1-\varepsilon}$ is smaller than n/k. Let X_t be the random variable counting the number of nodes of a specific color, say $h \in [k]$, at time t. For $T \leqslant c\,k \log n$ we thus have that

$$\mathbf{P}\left(X_T = n \,\middle|\, X_0 \leqslant \frac{n}{k} + \left(\frac{n}{k}\right)^{1-\varepsilon}\right) \leqslant \mathbf{P}\left(X_T \geqslant \frac{n}{k} + \left(1 + \frac{3}{k}\right)^T \left(\frac{n}{k}\right)^{1-\varepsilon} \,\middle|\, X_0 \leqslant \frac{n}{k} + \left(\frac{n}{k}\right)^{1-\varepsilon}\right) \quad (11)$$

[2]Notice that the inequality holds in particular for negative a as well

Given the initial condition $X_0 \leqslant \frac{n}{k} + \left(\frac{n}{k}\right)^{1-\varepsilon}$, if it holds that $X_T \geqslant \frac{n}{k} + \left(1+\frac{3}{k}\right)^T \left(\frac{n}{k}\right)^{1-\varepsilon}$, then a time step t with $0 \leqslant t \leqslant T-1$ must exist such that $X_t \leqslant n/k + b$ and $X_{t+1} \geqslant n/k + (1+3/k)b$ for some value b, with $k\sqrt{n\log n} \leqslant b \leqslant n/k$, thus

$$\mathbf{P}\left(X_T \geqslant \frac{n}{k} + \left(1+\frac{3}{k}\right)^T \left(\frac{n}{k}\right)^{1-\varepsilon} \,\Big|\, X_0 \leqslant \frac{n}{k} + \left(\frac{n}{k}\right)^{1-\varepsilon}\right) \leqslant$$

$$\begin{aligned}
\leqslant\ & \mathbf{P}\left(\exists 0 \leqslant t \leqslant T-1\ :\ X_{t+1} \geqslant \frac{n}{k} + \left(1+\frac{3}{k}\right) b \text{ and } X_t \leqslant \frac{n}{k} + b \,\Bigg|\, X_0 \leqslant \frac{n}{k} + \left(\frac{n}{k}\right)^{1-\varepsilon}\right) \\
& \text{(for some } b \text{ with } k\sqrt{n\log n} \leqslant b \leqslant n/k) \\
\leqslant\ & \sum_{t=0}^{T-1} \mathbf{P}\left(X_{t+1} \geqslant \frac{n}{k} + \left(1+\frac{3}{k}\right) b_t \text{ and } X_t \leqslant \frac{n}{k} + b_t \,\Bigg|\, X_0 \leqslant \frac{n}{k} + \left(\frac{n}{k}\right)^{1-\varepsilon}\right) \qquad (12) \\
& \text{(for some } b_0, \ldots, b_{T-1} \text{ with } k\sqrt{n\log n} \leqslant b_t \leqslant n/k \text{ for every } t = 0, \ldots, T-1)
\end{aligned}$$

Now observe that

$$\mathbf{P}\left(X_{t+1} \geqslant \frac{n}{k} + \left(1+\frac{3}{k}\right) b_t \text{ and } X_t \leqslant \frac{n}{k} + b_t \,\Bigg|\, X_0 \leqslant \frac{n}{k} + \left(\frac{n}{k}\right)^{1-\varepsilon}\right) =$$

$$\begin{aligned}
=\ & \sum_{a \leqslant b_t} \mathbf{P}\left(X_{t+1} \geqslant \frac{n}{k} + \left(1+\frac{3}{k}\right) b_t \text{ and } X_t = \frac{n}{k} + a \,\Bigg|\, X_0 \leqslant \frac{n}{k} + \left(\frac{n}{k}\right)^{1-\varepsilon}\right) \\
=\ & \sum_{a \leqslant b_t} \mathbf{P}\left(X_{t+1} \geqslant \frac{n}{k} + \left(1+\frac{3}{k}\right) b_t \,\Bigg|\, X_t = \frac{n}{k} + a \text{ and } X_0 \leqslant \frac{n}{k} + \left(\frac{n}{k}\right)^{1-\varepsilon}\right) \cdot \\
& \qquad \cdot \mathbf{P}\left(X_t = \frac{n}{k} + a \,\Bigg|\, X_0 \leqslant \frac{n}{k} + \left(\frac{n}{k}\right)^{1-\varepsilon}\right) \\
\leqslant\ & \frac{1}{n^2} \sum_{a \leqslant b_t} \mathbf{P}\left(X_t = \frac{n}{k} + a \,\Bigg|\, X_0 \leqslant \frac{n}{k} + \left(\frac{n}{k}\right)^{1-\varepsilon}\right) \leqslant \frac{1}{n^2} \qquad (13)
\end{aligned}$$

where in the last line we used Lemma 4.1.

By combining (11), (12), and (13) we get that, for every color $h \in [k]$, if the initial number of nodes colored h is $C_{0,h} \leqslant \frac{n}{k} + \left(\frac{n}{k}\right)^{1-\varepsilon}$ at any time $T \leqslant c\,k\log n$ the probability that all nodes are colored h is at most T/n^2. The probability that C_T is monochromatic is thus at most $(kT)/n^2 \leqslant n^{-\alpha}$ for some positive constant α. □

It may be worth noticing that what we actually prove in Theorem 4.2 is that $\Omega(k\log n)$ time steps are required in order to go from a configuration where the majority color has at most $n/k + (n/k)^{1-\varepsilon}$ nodes to a configuration where it has $2n/k$ colors.

4.2 A negative result for general 3-inputs dynamics

In order to prove that dynamics that differ from the majority ones do not solve the majority consensus, we first give some formal definitions of the dynamics we are considering.

Definition 4.3 ($\mathcal{D}_h(k)$ protocols) *An* h-dynamics *is a synchronous protocol where at each time step every node picks h random neighbors (including itself and with repetition) and recolors itself according to some deterministic rule that depends only on the colors it sees. Let $\mathcal{D}_h(k)$ be the class of h-dynamics and observe that a dynamics $\mathcal{P} \in \mathcal{D}_h$ can be specified by a function*

$$f : [k]^h \to [k]$$

such that $f(x_1, \ldots, x_h) \in \{x_1, \ldots, x_h\}$. Where $f(x_1, \ldots, x_h)$ is the color chosen by a node that sees the (ordered) sequence $(x_1, \ldots, x_h)$ of colors.

In the class $\mathcal{D}_3(k)$, there is a subset $\mathcal{M}^3$ of equivalent protocols called 3-majority dynamics having two key-properties described below: the clear-majority and the uniform one.

Definition 4.4 (clear-majority property) *Let $(x_1, x_2, x_3) \in [k]^3$ be a triple of colors. We say that (x_1, x_2, x_3) has a* clear majority *if at least two of the three entries have the same value. A dynamics $\mathcal{P} \in \mathcal{D}_3(k)$ has the* clear-majority *property if whenever its f sees a clear majority it returns the majority color.*

Given any 3-input dynamics function $f(x_1, x_2, x_3)$, for any triple of distinct colors $r, g, b \in [k]$, let $\Pi(r, g, b)$ be the subset of permutations of the colors r, g, b and define the following "counters":

$$\begin{aligned}
\delta_r &= |\{(z_1, z_2, z_3) \in \Pi(r, g, b), \ \ s.t. \ f(z_1, z_2, z_3) = r\}| \\
\delta_g &= |\{(z_1, z_2, z_3) \in \Pi(r, g, b), \ \ s.t. \ f(z_1, z_2, z_3) = g\}| \\
\delta_b &= |\{(z_1, z_2, z_3) \in \Pi(r, g, b), \ \ s.t. \ f(z_1, z_2, z_3) = b\}|
\end{aligned}$$

Observe that for any 3-inputs dynamics it must hold $\delta_g + \delta_r + \delta_b = 6$.

Definition 4.5 (uniform property) *A dynamics $\mathcal{P} \in \mathcal{D}_3(k)$ has the* uniform *property if, for any triple of distinct colors $r, g, b \in [k]$, it holds that $\delta_r = \delta_g = \delta_b$ $(= 2)$.*

Informally speaking, the clear-majority and the uniform properties provide a clean characterization of those dynamics that are good solvers for majority consensus. This fact is formalized in the next definitions and in the final theorem.

Definition 4.6 (3-majority dynamics) *A protocol $\mathcal{P} \in \mathcal{D}_3(k)$ belongs to the class $\mathcal{M}^3 \subset \mathcal{D}_3(k)$ of* 3-majority dynamics *if its function $f(x_1, x_2, x_3)$ has the clear-majority and the uniform properties.*

Definition 4.7 ((s, ε)-majority consensus solver) *We say that a protocol $\mathcal{P}$ is an (s, ε)-*solver *(for the majority consensus problem) if for every initial s-biased coloring $\overline{\mathrm{c}}$, when running $\mathcal{P}$, with probability at least $1 - \varepsilon$ there is a time step t by which all nodes gets the majority color of c.*

Let us observe that, by definition of h-dynamics, any monochromatic configuration is an absorbing state of the relative Markovian process. Moreover, the smaller s and ε the better an (s, ε)-solver is; in other words, if a dynamics is an (s, ε)-solver then it is also an (s', ε')-solver for every $s' \geqslant s$ and $\varepsilon' \geqslant \varepsilon$. In Section 3, we showed that any dynamics in $\mathcal{M}^3$ is a $\left(\Theta(\sqrt{\min\{2k, (n/\log n)^{1/3}\} n \log n}), \Theta(1/n)\right)$-solver in $\mathcal{D}_3$. We can now state the main result of this section.

Theorem 4.8 (properties for good solvers) *(a) If a protocol $\mathcal{P}$ is an $(n/4, 1/4)$-solver in $\mathcal{D}_3$ then its f must have the clear-majority property.*

(b) A constant $\eta > 0$ exists such that, if $\mathcal{P}$ is an $(\eta \cdot n, 1/4)$-solver, then its f must have the uniform property.

The above theorem also provides the clear reason why some dynamics can solve consensus but cannot solve majority consensus in the non-binary case. A relevant example is the *median* dynamics studied in [6]: it has the clear-majority property but not the uniform one.

For readability sake, we split the proof of the above theorem in two technical lemmas: in the first one, we show the first claim about clear majority while in the second lemma we show the second claim about the uniform property.

Lemma 4.9 (clear majority) *If a protocol $\mathcal{P} \in \mathcal{D}_3$ is an $(n/4, 1/4)$-solver, then it chooses the majority color every time there is a triple with a clear majority.*

Proof. For every triple of colors $(x_1, x_2, x_3) \in [k]^3$ that has a clear majority, let us define $\delta(x_1, x_2, x_3)$ to be 1 if protocol $\mathcal{P}$ behaves like the majority protocol over triple (x_1, x_2, x_3) and 0 otherwise. Consider an initial configuration with only two colors, say red (r) and blue (b), with c_r red nodes and $c_b = n - c_r$ blue nodes. Let us define Δ_r and Δ_b as follows

$$\begin{aligned}
\Delta_r &= \delta(r, r, b) + \delta(r, b, r) + \delta(b, r, r) \\
\Delta_b &= \delta(b, b, r) + \delta(b, r, b) + \delta(r, b, b)
\end{aligned}$$

We can write the probability that a node chooses color red as

$$\begin{aligned}
p(r) &= \left(\frac{c_r}{n}\right)^3 + \left(\frac{c_r}{n}\right)^2 \frac{c_b}{n} \cdot \Delta_r + \left(\frac{c_b}{n}\right)^2 \frac{c_r}{n} (3 - \Delta_b) \\
&= \frac{c_r}{n^3} \left(c_r^2 + c_b \left(c_r \Delta_r - c_b \Delta_b\right) + 3c_b^2\right) \qquad (14)
\end{aligned}$$

Observe that for a majority protocol we have that $\Delta_r = \Delta_b = 3$. In what follows we show that if this is not the case then there are configurations where the majority color does not increase in expectation. We distinguish two cases, case $\Delta_r \neq \Delta_b$ and case $\Delta_r = \Delta_b$.

<u>Case $\Delta_r \neq \Delta_b$</u>: Suppose w.l.o.g. that $\Delta_r < \Delta_b$, and observe that since they have integer values it means $\Delta_r \leqslant \Delta_b - 1$. Now we show that, if we start from a coloring where the red color has the majority of nodes, the number of red nodes decreases in expectation. By using $\Delta_r \leqslant \Delta_b - 1$ in (14) we get

$$p(r) \leqslant \frac{c_r}{n^3} \left(c_r^2 + c_b(c_r - c_b)\Delta_b - c_r c_b + 3c_b^2\right) \qquad (15)$$

If the majority of nodes is red then $c_r - c_b$ is positive, and since Δ_b can be at most 3 from (15) we get

$$p(r) \leqslant \frac{c_r}{n^3} \left(c_r^2 + 2c_r c_b\right) \qquad (16)$$

Finally, if we put $c_r = n/2 + s$ and $c_b = n/2 - s$, for some positive s, in (16), we get that

$$p(r) \leqslant \frac{c_r}{n^3} \left(\frac{3}{4}n^2 + (n - s)s\right) \leqslant \frac{c_r}{n} \qquad (17)$$

Case $\Delta_r = \Delta_b$: When $\Delta_r = \Delta_b$, observe that if the protocol is not a majority protocol then it must be $\Delta_r = \Delta_b \leqslant 2$. Hence, if we start again from a configuration where $c_r \geqslant c_b$, from (14) we get that

$$p(r) \leqslant \frac{c_r}{n^3}\left(c_r^2 + 2c_b(c_r - c_b) + 3c_b^2\right) = \frac{c_r}{n} \tag{18}$$

In both cases, for any protocol $\mathcal{P}$ that does not behave like a majority protocol on triples with a clear majority, if we name X_t the random variable indicating the number of red nodes at time t, from (17) and (18) we get that $\mathbf{E}\left[X_{t+1} \mid X_t\right] \leqslant X_t$, hence X_t is a supermartingale. Now let τ be the random variable indicating the first time the chain hits one of the two absorbing states, i.e.

$$\tau = \inf\{t \in \mathbb{N} \,:\, X_t \in \{0, n\}\}$$

Since $\mathbf{P}\left(\tau < \infty\right) = 1$ and all X_t's have values bounded between 0 and n, from the martingale stopping theorem[3] we get that $\mathbf{E}\left[X_\tau\right] \leqslant \mathbf{E}\left[X_0\right]$. If we start from a configuration that is $n/4$-unbalanced in favor of the red color, we have that $X_0 = n/2 + n/8$, and if we call ε is the probability that the process ends up with all blue nodes we have that $\mathbf{E}\left[X_\tau\right] = (1-\varepsilon)n$. Hence it must be $(1-\varepsilon)n \leqslant n/2 + n/8$ and the probability to end up with all blue nodes is $\varepsilon \geqslant 5/8 > 1/4$. Thus the protocol is not a $(n/4, 1/4)$-solver. □

Lemma 4.10 (uniform property) *A constant $\eta > 0$ exists such that, if $\mathcal{P}$ is an $(\eta n, 1/4)$-solver, then its f must have the uniform property.*

Proof. Thanks to the previous lemma, we can assume that $f$has the clear-majority property but a triple (r, g, b) exists such that $\delta_r < \max\{\delta_g, \delta_b\}$. Let's start the process with the following initial configuration having only the above 3 colors and then show that the process w.h.p. will not converge to the majority color r.

$$\overline{c} = (c_r, c_g, c_b), \quad \text{where} \quad c_r = \frac{n}{3} + s,\ c_g = n/3,\ c_b = \frac{n}{3} - s \quad \text{with } s = \Theta(\sqrt{n \log n})$$

We consider the "hardest" case where $\delta_r = 1$: the case $\delta_r = 0$ is simpler since in this case, no matter how the other $\delta' s$ are distributed, it is easy to see that the r.v. c_r will decrease exponentially to 0 starting from the above configuration.

- Case $\delta_r = 1$, $\delta_g = 3$, and $\delta_b = 2$ (and color-symmetric cases). Starting from the above initial configuration, we can compute the probability $p(r) = \mathbf{P}\left(X_v = r \mid C = \overline{c}\right)$ that a node gets the color r.

$$\begin{aligned} p(r) &= \left(\frac{c_r}{n}\right)^3 + 3\left(\frac{c_r}{n}\right)^2 \frac{n - c_r}{n} + \frac{c_r c_g c_b}{n^3} \\ &= \frac{n + 3s}{3n^3}\left(\left(\frac{n}{3} + s\right)^2 + 3\left(\frac{n}{3} + s\right)\left(\frac{2}{3}n - s\right) + \left(\frac{n}{3}\right)\left(\frac{n}{3} - s\right)\right) \end{aligned}$$

After some easy calculations, we get

$$p(r) = \frac{8}{27}\left(1 + O\left(\frac{s}{n}\right)\right)$$

As for $p(g)$, by similar calculations, we obtain the following bound

[3]See e.g. Chapter 17 in [14] for a summary of martingales and related results

$$p(g) = \frac{10}{27}\left(1 - O\left(\frac{s^2}{n^2}\right)\right)$$

From the above two equations, we immediately have the following bounds on the expectation of the r.v.'s X^r and X^g counting the nodes colored with r and g, respectively (at the next time step).

$$\mathbb{E}\left[X^r \mid C = \overline{c}\right] \leqslant \frac{8}{27}n + O(s) \text{ and } \mathbb{E}\left[X^g \mid C = \overline{c}\right] \geqslant \frac{10}{27}n - O\left(\frac{s^2}{n}\right)$$

By a standard application of Chernoff's Bound, we can prove that, if $s \leqslant \eta n$ for a sufficiently small $\eta > 0$, the initial value c_r will w.h.p. decreases by a constant factor, going much below the new majority c_g. Then, by applying iteratively the above reasoning we get that the process w.h.p. will not converge to r.

- Case $\delta_r = 1$, $\delta_g = 4$, and $\delta_b = 1$ (and color-symmetric cases). In this case it is even simpler to show that w.h.p., starting from the same initial configuration considered in the previous case, the process will not converge to color r. □

4.3 A lower bound for h-majority

In Subsection 4.1, we have shown that the 3-majority protocol takes $\Theta(k \log n)$ time steps w.h.p. to converge in the worst case. A natural question is whether by using the h-majority protocol, with h slightly larger than 3, it is possible to significantly speed-up the process. We prove that this is not the case.

Let us consider a set of n nodes, each node colored with one out of k colors. The h-majority protocol runs as follows:

> *At every time step, every node picks h nodes uniformly at random (including itself and with repetitions) and recolors itself according to the majority of the colors it sees (breaking ties u.a.r.)*

Let $j \in [k]$ be an arbitrary color, in the next lemma we prove that, if the number of j-colored nodes is smaller than $2n/k$ and if $k/h = \mathcal{O}(n^{(1-\varepsilon)/4})$, then the probability that the number of j-nodes increases by a factor $(1 + \Theta(h^2/k))$ is exponentially small.

Lemma 4.11 *Let $j \in [k]$ be a color and let X_t be the random variable counting the number of j-colored nodes at time t. If $k/h = \mathcal{O}(n^{(1-\varepsilon)/4})$, then for every $(n/k) \leqslant a \leqslant 2(n/k)$ it holds that*

$$\mathbf{P}\left(X_{t+1} \geqslant \left(1 + \frac{h^2}{k}\right) a \mid X_t = a\right) \leqslant e^{-\Theta(n^{\varepsilon})}$$

Proof. Consider a specific node, say $u \in [n]$, let N_j be the number of j-colored nodes picked by u during the sampling stage of the t-th time step and let Y be the indicator random variable of the event that node u chooses color j at time step $t+1$. We give an upper bound on the probability of the event $Y = 1$ by conditioning it on $N_j = 1$ and $N_j \geqslant 2$ (observe that if $N_j = 0$ node u cannot choose j as its color at the next time step)

$$\mathbf{P}\left(Y_u = 1\right) \leqslant \mathbf{P}\left(Y_u = 1 \mid N_j = 1\right)\mathbf{P}\left(N_j = 1\right) + \mathbf{P}\left(N_j \geqslant 2\right) \tag{19}$$

Now observe that

- $\mathbf{P}\left(Y_u = 1 \,|\, N_j(u) = 1\right) \leqslant 1/h$, since it is exactly $1/h$ if all other sampled nodes have distinct colors and it is 0 otherwise;
- $\mathbf{P}\left(N_j = 1\right) \leqslant h\frac{a}{n}$, since it can be bounded by the probability that at least one of the h samples gives color j;
- $\mathbf{P}\left(N_j \geqslant 2\right) \leqslant \binom{h}{2}\frac{a^2}{n^2}$, since it is the probability that a pair of sampled nodes exist with the same color j.

Hence in (19) we have that

$$\mathbf{P}\left(Y = 1\right) \leqslant \frac{a}{n} + \frac{h^2}{2} \cdot \frac{a^2}{n^2}$$

Thus, for the expected number of j-colored nodes at the next time step we get

$$\mathbf{E}\left[X_{t+1} \,|\, X_t = a\right] \leqslant a + \frac{h^2}{2n}a^2 = a\left(1 + \frac{h^2}{2n}a\right) \leqslant a\left(1 + \frac{h^2}{k}\right)$$

where in the last inequality we used the hypothesis $a \leqslant 2(n/k)$.

Since X_{t+1} is a sum of n independent Bernoulli random variables, from Chernoff bound (Lemma A.2 with $\lambda = ah^2/k$) we finally get

$$\mathbf{P}\left(X_{t+1} \geqslant a\left(1 + 2\frac{h^2}{k}\right) \,|\, X_t = a\right) \leqslant \exp\left(-\frac{2(ah^2/k)^2}{n}\right) \leqslant \exp\left(-\Omega(n^{\varepsilon})\right)$$

where in the last inequality we used $a \geqslant n/k$ and $k/h = \mathcal{O}(n^{(1-\varepsilon)/4})$. □

By adopting a similar argument to that used for proving Theorem 4.2, we can prove a lower bound $\Omega(k/h^2)$ on the completion time of the h-majority.

Theorem 4.12 *Let $\mathbf{C}_t$ be the random variable indicating the coloring at time t according to the h-majority protocol and let $\tau = \inf\{t \in \mathbb{N} : \mathbf{C}_t \text{ is monochromatic}\}$. If the initial coloring is $\mathbf{C}_0 = (c_1, \ldots, c_k)$ with $\max\{c_j : j = 1, \ldots, k\} \leqslant \frac{3}{2} \cdot \frac{n}{k}$ then $\tau = \Omega(k/h^2)$ w.h.p.*

Proof. Since in the initial coloring the majority color has $a \leqslant (3/2)(n/k)$ nodes, from Lemma 4.11 it follows that the number of nodes with the majoritycolor increases at a rate smaller than $(1 + 2h^2/k)$ with probability exponentially close to 1. Hence it follows a recursive relation of the form $X_{t+1} \leqslant \left(1 + 2h^2/k\right) X_t$, which gives

$$X_t \leqslant \left(1 + 2h^2/k\right)^t X_0 \leqslant \left(1 + 2h^2/k\right)^t \frac{3}{2} \cdot \frac{n}{k}$$

We thus have that

$$(3/2)\left(1 + \frac{2h^2}{k}\right)^t \leqslant 2 \;\text{ for }\; t \leqslant \frac{k}{h^2}\log(4/3)$$

□

5 Open Questions

Several interesting issues are still open on the majority consensus problem even when simple distributed models are considered. A first one is whether an updating rule exists that, by using some (small, i.e. $o(\log n)$) extra memory, can guarantee majority consensus in polylogarithmic time for any value of k. We suspect that this might be possible but at the "cost" of loosing the self-stabilizing property of the simpler dynamics such as 3-majority (remind the discussion in the Introduction): in other words, the use of some extra memory bits (that somewhat store "local" statistics about colors) might make dynamics much more susceptible to adversaries that can corrupt such bits at every round.

A more specific question about our simple distributed model is to explore what happens when the initial bias s is smaller than the lower bounds assumed in our analysis. Notice that when k is polylogarithmic, the required bias is only a polylogarithmic factor larger than the standard deviation $\Omega(\sqrt{n})$ which is a lower bound for the initial bias to converge (w.h.p.) to the majority color. As for larger k, we cannot claim any stronger bound on the required bias, however, in Appendix A.3, we show that there are initial configurations with bias $s = O(\sqrt{kn})$ for which the initial bias *decreases* in a single round with constant probability. This result implies that, when the initial bias s is "slightly" smaller than that assumed in our upper bound ($s \geqslant c\sqrt{\min\{2k, (n/\log n)^{1/3}\}\, n \log n}$), the process may be *non-monotone* w.r.t. the bias function $s(t)$. On the contrary, the fact that $s(t)$ is an increasing function played a key-role in the proof of our upper bound. So, under such a weaker assumption, if any upper bound similar to ours might be proved then a much more complex argument (departing from ours) seems to be necessary.

References

[1] M. A. Abdullah and M. Draief. Majority consensus on random graphs of a given degree sequence. *arXiv preprint arXiv:1209.5025*, 2012.

[2] D. Angluin, J. Aspnes, and D. Eisenstat. A simple population protocol for fast robust approximate majority. In *DISC*, LNCS, 2007.

[3] A. Babaee and M. Draief. Distributed multivalued consensus. In *Computer and Information Sciences III*, pages 271–279. Springer, 2013.

[4] F. Bénézit, P. Thiran, and M. Vetterli. Interval consensus: from quantized gossip to voting. In *ICASSP*, pages 3661–3664. IEEE, 2009.

[5] A. Clementi, M. Di Ianni, G. Gambosi, E. Natale, and R. Silvestri. Distributed community detection in dynamic graphs. In *SIROCCO*. Springer, 2013.

[6] B. Doerr, L. A. Goldberg, L. Minder, T. Sauerwald, and C. Scheideler. Stabilizing consensus with the power of two choices. In *SPAA*. ACM, 2011.

[7] M. Draief and M. Vojnovic. Convergence speed of binary interval consensus. *SIAM J. on Control and Optimization*, 50(3):1087–1109, 2012.

[8] D. Dubhashi and D. Ranjan. Balls and bins: A study in negative dependence. *Random Structures and Algorithms*, 13(2):99–124, 1998.

[9] D. P. Dubhashi and A. Panconesi. *Concentration of measure for the analysis of randomized algorithms*. Cambridge University Press, 2009.

[10] D. Easley and J. Kleinberg. *Networks, Crowds, and Markets*. Cambridge Univ Press, 2010.

[11] S. Greenberg and M. Mohri. Tight lower bound on the probability of a binomial exceeding its expectation. *arXiv preprint arXiv:1306.1433*, 2013.

[12] M. Kearns and J. Tan. Biased voting and the democratic primary problem. In *WINE*, pages 639–652. Springer, 2008.

[13] M. W. S. Land and R. K. Belew. No two-state ca for density classification exists. *Phys. Rev. Letters*, 74(25):5148–5150, 1995.

[14] D. Levin, Y. Peres, and E. L. Wilmer. *Markov Chains and Mixing Times*. AMS, 2008.

[15] E. Mossel, J. Neeman, and O. Tamuz. Majority dynamics and aggregation of information in social networks. *Autonomous Agents and Multi-Agent Systems*, pages 1–22, 2012.

[16] E. Mossel and G. Schoenebeck. Reaching consensus on social networks. In *ICS*, pages 214–229, 2010.

[17] N. Mousavi. How tight is chernoff bound? url: https://ece.uwaterloo.ca/ nmousavi/Papers/Chernoff-Tightness.pdf.

[18] D. Peleg. Local majorities, coalitions and monopolies in graphs: a review. *Theor. Comput. Sci.*, 282(2), 2002.

[19] E. Perron, D. Vasudevan, and M. Vojnovic. Using three states for binary consensus on complete graphs. In *INFOCOM*, pages 2527–2535. IEEE, 2009.

A Appendix

A.1 Inequalities

Lemma A.1 (Chernoff bounds - multiplicative form) *Let $Y_1, Y_2, \ldots$ be independent 0-1 random variables. Let $Y = \sum_i Y_i$ and let $\mu = \mathbb{E}[Y]$. Then,*

1. *for any $0 < \delta \leqslant 4$,*
$$\mathbf{P}(Y > (1+\delta)\mu) < e^{-\frac{\delta^2 \mu}{4}} \tag{20}$$

2. *for any $\delta \geqslant 4$,*
$$\mathbf{P}(Y > (1+\delta)\mu) < e^{-\delta\mu} \tag{21}$$

Lemma A.2 (Chernoff bound - additive form) *Let $X = \sum_{i=1}^{n} X_i$ where X_i's are independent Bernoulli random variables and let $\mu = \mathbf{E}[X]$. Then for every $\lambda > 0$ it holds that*

$$\mathbf{P}(X \geqslant \mu + \lambda) \leqslant e^{-2\lambda^2/n}$$

Lemma A.3 (Jensen inequality) *Let $\phi : \mathbb{R} \to \mathbb{R}$ be a convex function and $x_1, \ldots x_k \in \mathbb{R}$ be k real numbers, then*

$$\phi\left(\frac{1}{k}\sum_{i=1}^{k} x_i\right) \leqslant \frac{1}{k}\sum_{i=1}^{k} \phi(x_i)$$

Lemma A.4 (Bernstein inequality [9]) *Let the random variables $X_1, \ldots, X_n$ be independent with $X_i - \mathbb{E}[X_i] \leqslant b$ for each $i \in [n]$. Let $X = \sum_i X_i$ and let $\sigma^2 = \sum_i \sigma_i^2$ be the variance of X. Then, for any $\lambda > 0$,*

$$\mathbf{P}(X > \mathbb{E}[X] + \lambda) \leqslant \exp\left(-\frac{\lambda^2}{2\sigma^2(1 + b\lambda/3\sigma^2)}\right)$$

A.2 Technical lemma about Markov chains

Lemma A.5 *Let $\mathcal{M}$ be any finite Markov chain and let S be its set of states. Let X_t be the random variable representing the state of the chain at time t. If $A_1, \ldots, A_T$ are such that $S \supseteq A_1 \supseteq A_2 \supseteq \cdots \supseteq A_T$ and, for any $i = 1, \ldots, T$,*

$$\mathbf{P}(X_t \in A_i \mid X_{t-1} \in A_i) \geqslant 1 - \epsilon \quad \text{and, for } i < T, \quad \mathbf{P}(X_t \in A_{i+1} \mid X_{t-1} \in A_i) \geqslant 1 - \nu$$

where $0 \leqslant \epsilon \leqslant \nu < 1$, then, for any integer $\ell \geqslant 1$,

$$\mathbf{P}(X_{\ell T} \in A_T \mid X_0 \in A_1) \geqslant 1 - T(\ell\,\epsilon + \nu^\ell)$$

Proof. Firstly we prove the following

Claim 1 *If A and B are two sets of states such that $A \supseteq B$,*

$\mathbf{P}(X_t \in A \mid X_{t-1} \in A) \geqslant 1 - \epsilon$, $\quad \mathbf{P}(X_t \in B \mid X_{t-1} \in B) \geqslant 1 - \epsilon$, *and*
$\mathbf{P}(X_t \in B \mid X_{t-1} \in A) \geqslant 1 - \nu$

Then, for any integer $\ell \geqslant 1$*,*

$$\mathbf{P}\left(X_\ell \in B \mid X_0 \in A\right) \geqslant (1-\epsilon)^\ell - (\nu-\epsilon)^\ell$$

Proof of claim. The proof is by induction on ℓ. If $\ell = 1$, it is immediate from the hypotheses. For the sake of brevity, in the sequel we omit the conditioning "$X_0 \in A$". Let $\ell \geqslant 2$. It holds that

$$\begin{aligned}
\mathbf{P}\left(X_\ell \in B\right) &= \mathbf{P}\left(X_\ell \in B \mid X_{\ell-1} \in B\right)\mathbf{P}\left(X_{\ell-1} \in B\right) + \\
&\quad + \mathbf{P}\left(X_\ell \in B \mid X_{\ell-1} \in A-B\right)\mathbf{P}\left(X_{\ell-1} \in A-B\right) + \\
&\quad + \mathbf{P}\left(X_\ell \in B \mid X_{\ell-1} \notin B\right)\mathbf{P}\left(X_{\ell-1} \notin B\right) \\
&\geqslant \mathbf{P}\left(X_\ell \in B \mid X_{\ell-1} \in B\right)\mathbf{P}\left(X_{\ell-1} \in B\right) + \\
&\quad + \mathbf{P}\left(X_\ell \in B \mid X_{\ell-1} \in A-B\right)\mathbf{P}\left(X_{\ell-1} \in A-B\right) \\
&\geqslant (1-\epsilon)\mathbf{P}\left(X_{\ell-1} \in B\right) + (1-\nu)\mathbf{P}\left(X_{\ell-1} \in A-B\right) \qquad (22)
\end{aligned}$$

Observe that

$$\begin{aligned}
\mathbf{P}\left(X_{\ell-1} \in A-B\right) &= 1 - \mathbf{P}\left(X_{\ell-1} \in B\right) - \mathbf{P}\left(X_{\ell-1} \notin A\right) \\
&= \mathbf{P}\left(X_{\ell-1} \in A\right) - \mathbf{P}\left(X_{\ell-1} \in B\right) \\
&\geqslant \mathbf{P}\left(X_0 \in A\right)\prod_{t=1}^{\ell-1}\mathbf{P}\left(X_t \in A \mid X_{t-1} \in A\right) - \mathbf{P}\left(X_{\ell-1} \in B\right) \\
&\geqslant (1-\epsilon)^{\ell-1} - \mathbf{P}\left(X_{\ell-1} \in B\right)
\end{aligned}$$

By taking into account the above inequality in Ineq. 22, we obtain

$$\begin{aligned}
\mathbf{P}\left(X_\ell \in B\right) &\geqslant (1-\epsilon)\mathbf{P}\left(X_{\ell-1} \in B\right) + (1-\nu)((1-\epsilon)^{\ell-1} - \mathbf{P}\left(X_{\ell-1} \in B\right)) \\
&= (\nu-\epsilon)\mathbf{P}\left(X_{\ell-1} \in B\right) + (1-\nu)(1-\epsilon)^{\ell-1} \\
&\geqslant (\nu-\epsilon)((1-\epsilon)^{\ell-1} - (\nu-\epsilon)^{\ell-1}) + (1-\nu)(1-\epsilon)^{\ell-1} \\
&\qquad \text{(by the inductive hypothesis)} \\
&= (1-\epsilon)^\ell - (\nu-\epsilon)^\ell
\end{aligned}$$

□

It holds that

$$\begin{aligned}
\mathbf{P}\left(X_{\ell T} \in A_T \mid X_0 \in A_1\right) &\geqslant \mathbf{P}\left(X_\ell \in A_1 \mid X_0 \in A_1\right) \prod_{i=2}^{T} \mathbf{P}\left(X_{\ell i} \in A_i \mid X_{\ell(i-1)} \in A_{i-1}\right) \\
&\geqslant (1-\epsilon)^\ell \prod_{i=2}^{T}((1-\epsilon)^\ell - (\nu-\epsilon)^\ell) \quad \text{(from Claim 1)} \\
&= (1-\epsilon)^\ell((1-\epsilon)^\ell - (\nu-\epsilon)^\ell)^{T-1} \\
&= (1-\epsilon)^{\ell T}\left(1-\left(\frac{\nu-\epsilon}{1-\epsilon}\right)^\ell\right)^{T-1} \\
&\geqslant (1-\epsilon)^{\ell T}\left(1-T\left(\frac{\nu-\epsilon}{1-\epsilon}\right)^\ell\right) \\
&= (1-\epsilon)^{\ell T} - T(1-\epsilon)^{\ell T}\left(\frac{\nu-\epsilon}{1-\epsilon}\right)^\ell \\
&\geqslant 1-\ell T\epsilon - T\frac{(1-\epsilon)^{\ell T}}{(1-\epsilon)^\ell}(\nu-\epsilon)^\ell \\
&\geqslant 1-T(\ell\epsilon+\nu^\ell)
\end{aligned}$$

□

A.3 On the initial bias

In this section, we show that there are initial configurations with bias $s = O(\sqrt{kn})$ for which the initial bias decreases in a single round with constant probability.

We consider an initial configuration in which we have $k \geqslant 3$ colors, for k larger than a suitable constant. Let $s = O(\sqrt{kn})$ and $C = \frac{n-s}{k}$. We assume that the majority color is m, with $c_m = C + s$, while $C_j = C$, for $j \neq m$. We thus have $k - 1 = \frac{n-C-s}{C}$ (we neglect integer parts for the sake of the analysis). We also assume that $s \leqslant C$ (and in any case $s \leq \frac{n}{k}$) and let C'_j be the random variable indicating the number of j-colored nodes at the next time step. For any $j \neq m$, we want to prove that $C'_m - C'_j < s$ with constant probability. From Lemma 2.1, we easily get the following equations

$$\mathbb{E}\left[C'_m\right] = C + s + \frac{C^2}{n} + \frac{2Cs+s^2}{n} - \frac{C+s}{n^2}\gamma \quad \text{and} \quad \mathbb{E}\left[C'_j\right] = C + \frac{C^2}{n} - \frac{C}{n^2}\gamma$$
$$\text{where} \quad \gamma = \sum_h C_h^2 = nC + Cs + s^2$$

We will also use the following "reverse"-Chernoff bound [17, Theorem 2]

Theorem A.6 *Let X be the sum of m independent Bernoulli variables with probability $p \leq \frac{1}{4}$ and let $\mu = pm$. Then, for any $t > 0$:*

$$\mathbf{P}\left(X - \mu > t\right) \geq \frac{1}{4}e^{-\frac{2t^2}{\mu}}.$$

We can now state the result of this section.

Lemma A.7 *For any color $j \neq m$ and for an initial bias $s \leq \frac{\sqrt{kn}}{36}$, it holds that*

$$\mathbf{P}\left(C'_m - C'_j < s\right) \geq \frac{1}{16e}.$$

Proof. We first show that $\mathbb{E}\left[C'_m\right] - \mathbb{E}\left[C'_j\right] \leqslant s + \frac{3Cs}{n}$, then we observe that with constant probablity $\mathbb{E}\left[C'_m\right]$ is not above its average, and, finally, we prove that $C'_j > \mathbb{E}\left[C'_j\right] + \frac{3Cs}{n}$ with constant probability, whenever $s \leq \frac{\sqrt{kn}}{36}$. This is enough to prove the lemma.
Proof. By definition of γ we get

$$\mathbb{E}\left[C'_m\right] = C + s + \frac{C^2}{n} + \frac{2Cs + s^2}{n} - \frac{C+s}{n^2}(nC + Cs + s^2) = C + s + \frac{Cs}{n} + \frac{s^2}{n} - \frac{s}{n^2}(C+s)^2$$
$$\leqslant C + s + \frac{2Cs}{n} - \frac{s}{n^2}(C+s)^2 < C + s + \frac{2Cs}{n},$$

where the third inequality follows by assuming $s \leqslant C$. Analogously we have:

$$\mathbb{E}\left[C'_j\right] = C + \frac{C^2}{n} - \frac{C}{n^2}(nC + Cs + s^2) = C - \frac{Cs}{n} \cdot \frac{C+s}{n} \geqslant C - \frac{2C^2 s}{n^2}, \tag{23}$$

where to derive the last inequality we again use $s \leqslant C$. We also have:

$$\mathbb{E}\left[C'_m\right] - \mathbb{E}\left[C'_j\right] = s + \frac{2Cs + s^2}{n} - \frac{s\gamma}{n^2} \leqslant s + \frac{3Cs}{n},$$

where we again assume that $s \leqslant C$. Next, we note that

$$\mathbf{P}\left(C'_m - C'_j < s\right) \geq \mathbf{P}\left(C'_m < \mathbb{E}\left[C'_j\right] + s\frac{4Cs}{n} \wedge C'_j \geq \mathbb{E}\left[C'_j\right] + \frac{4Cs}{n}\right).$$

If we consider the events $A = \left(C'_m < \mathbb{E}\left[C'_j\right] + s\frac{4Cs}{n}\right)$ and $B = \left(C'_j \geq \mathbb{E}\left[C'_j\right] + \frac{4Cs}{n}\right)$ then it is not hard to show the following

Fact 1

$$\mathbf{P}\left(A \wedge B\right) \geq \mathbf{P}\left(A\right)\mathbf{P}\left(B\right).$$

Proof. We have:

$$\mathbf{P}\left(A \wedge B\right) = \mathbf{P}\left(A \mid B\right)\mathbf{P}\left(B\right) = (1 - \mathbf{P}\left(\neg A \mid B\right))\mathbf{P}\left(B\right) = \mathbf{P}\left(B\right) - \mathbf{P}\left(\neg A \wedge B\right)$$
$$\geqslant \mathbf{P}\left(B\right) - \mathbf{P}\left(\neg A\right)\mathbf{P}\left(B\right) = \mathbf{P}\left(B\right) - (1 - \mathbf{P}\left(A\right))\mathbf{P}\left(B\right) = \mathbf{P}\left(A\right)\mathbf{P}\left(B\right),$$

where the fourth inequality follows from [8, Proposition 3, claim (-OD)]. In particular, $(\neg A)$ and (B) are the events that the numbers of balls thrown independently at random into two distinct bins both exceed some given thresholds. □

Fact 2

$$\mathbf{P}\left(C'_m \leq \mathbb{E}\left[C'_m\right]\right) \geq \frac{1}{4}.$$

Proof. Set $\tilde{C} = n - C'_m$. Clearly, $\tilde{C}$ is distributed as $B(n,p)$, where $p = 1 - p_m$, with p_m the probability that the generic node recolors itself with color m. Clearly, $p > \frac{1}{n}$ as long as the number of colors is not too large (in the order of n). Then we have:

$$\mathbf{P}\left(C'_m \leq \mathbb{E}\left[C'_m\right]\right) = \mathbf{P}\left(\tilde{C} \geq \mathbb{E}\left[\tilde{C}\right]\right) \geq \frac{1}{4},$$

where the second inequality follows from [11, Theorem 1]. $\square$

By applying Theorem A.6 to C'_j we have:

$$\mathbf{P}\left(C'_j > \mathbb{E}\left[C'_j\right] + \frac{3Cs}{n}\right) \geq \frac{1}{4} e^{-\frac{18C^2 s^2}{n^2 \mathbb{E}\left[C'_j\right]}} \geq \frac{1}{4} e^{-\frac{18Cs^2}{n^2 - 2Cs}}$$

where the second inequality follows from Inequality (23).

We apply Theorem A.6 to C'_j and we have:

$$\mathbf{P}\left(C'_j > \mathbb{E}\left[C'_j\right] + \frac{3Cs}{n}\right) \geq \frac{1}{4} e^{-\frac{18C^2 s^2}{n^2 \mathbb{E}\left[C'_j\right]}} \geq \frac{1}{4} e^{-\frac{18Cs^2}{n^2 - 2Cs}} \geqslant \frac{1}{4e} \tag{24}$$

where the second inequality follows from Inequality (23) and the third one holds since $s \leqslant \sqrt{kn}/36$ and $C \leqslant n/k$. Finally, from Fact 2 and Equation 24, we get the claim. $\square$